PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/31, C07K 14/22, A61K 39/02, 39/095, C07K 19/00, 16/12, C12Q 1/68, 1/02, A01N 33/53

A2

(11) International Publication Number: **WO 98/06851**

(43) International Publication Date: 19 February 1998 (19.02.98)

(21) International Application Number: PCT/US97/14221

(22) International Filing Date: 12 August 1997 (12.08.97)

(30) Priority Data:
60/023,832 12 August 1996 (12.08.96) US

(60) Parent Application or Grant
(63) Related by Continuation
US 60/023,832 (CIP)
Filed on 12 August 1996 (12.08.96)

(71) Applicant *(for all designated States except US)*: BOARD OF REGENTS, THE UNIVERSITY OF TEXAS SYSTEM [US/US]; 201 West 7th Street, Austin, TX 78701 (US).

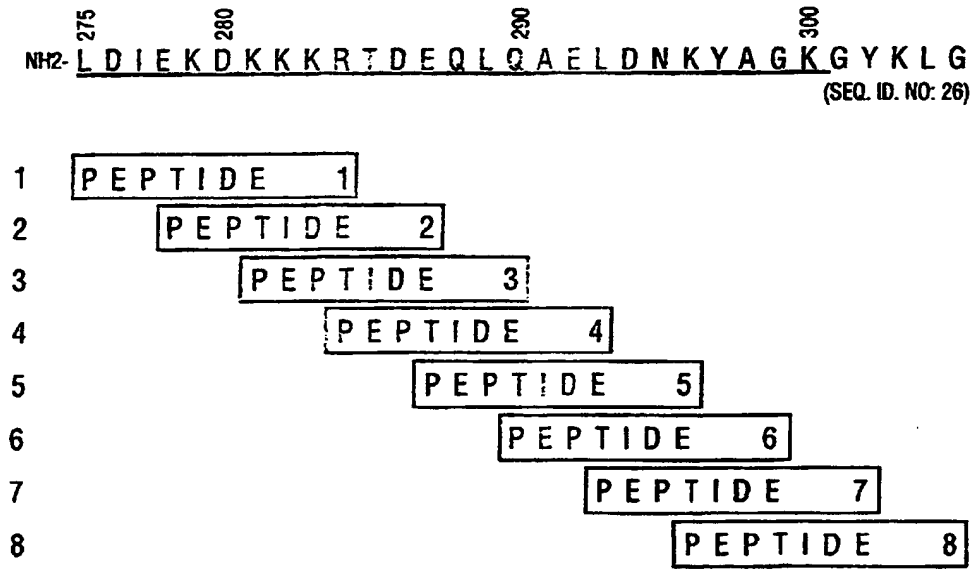
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: HANSEN, Eric, J. [US/US]; 2404 Chamberlain Drive, Plano, TX 75023 (US). COPE, Leslie, D. [US/US]; 3244 North Galloway #118, Mesquite, TX 75150 (US). AEBI, Christoph [CH/CH]; Gaselstrasse 165, CH-3144 Gasel (CH).

(74) Agent: HIGHLANDER, Steven, L.; Arnold, White & Durkee, P.O. Box 4433, Houston, TX 77210 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: DEFINING EPITOPES OF THE OUTER MEMBRANE PROTEIN CopB OF *MORAXELLA CATARRHALIS*



(57) Abstract

The present invention involves an analysis of the CopB outer membrane protein of *M. catarrhalis*. Four CopB antigens now have been sequenced and their predicted amino acid sequences compared. Regions of conservation and non-conservation have been identified in this antigen, including one non-conserved region that represents an antibody binding region from the strain 035E. A single amino acid change (N to D) in this epitope, at residue 295, abolishes reactivity of the antibody 10F3 with CopB. Peptides which only contain residues of this region that are C-terminal to residue 295 do retain reactivity but their reactivity is less than the maximal reactivity achieved in the presence of an asparagine at position 295. Compositions and both diagnostic and immunoprophylactic methods are disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

DESCRIPTION

DEFINING EPITOPES OF THE OUTER MEMBRANE PROTEIN CopB OF *MORAXELLA CATARRHALIS*

BACKGROUND OF THE INVENTION

This application is related to U.S. Serial No. 08/025,363, which is hereby incorporated by reference. The government owns rights in the present invention pursuant to grant number AI-36344 from the National Institutes of Health.

1. Field of the Invention

The present invention relates generally to the fields of microbiology and clinical bacteriology. In particular, sequences of the CopB outer membrane protein from various strains of *Moraxella catarrhalis* provide useful targets in immunodiagnosis and immunoprophylaxis.

2. Description of Related Art

It was previously thought that *M. catarrhalis,* previously known as *Branhamella catarrhalis* or *Neisseria catarrhalis* was a harmless saprophyte of the upper respiratory tract (Catlin, 1990; Berk, 1990). However, during the previous decade, it has been determined that this organism is an important human pathogen. Indeed, it has been established that this Gram-negative diplococcus is the cause of a number of human infections (Murphy, 1989).

M. catarrhalis is now known to be the third most common cause of both acute and chronic otitis media in children and accounts for up to 20% of cases (Aspin *et al.*, 1994; Del Beccaro *et al.*, 1992; Gan *et al.*, 1991; Catlin, 1990; Faden *et al.*, 1990;1991; Marchant, 1990), the most common disease for which infants and children receive health care according to the 1989 Consensus Report. This organism also causes acute maxillary sinusitis, generalized infections of the lower respiratory tract (Murphy and Loeb, 1989),

and is an important cause of bronchopulmonary infections in patients with underlying chronic lung disease and, less frequently, of systemic infections in immunocompromised patients (Melendez and Johnson, 1990; Sarubbi *et al.*, 1990; Schonheyder and Ejlertsen, 1989; Wright and Wallace, 1989).

In a recent report, *M. catarrhalis* DNA could be detected by polymerase chain reaction (PCR™) in middle ear fluid from 46% of patients with chronic otitis media with effusion (Post *et al.*, 1995). In adults, *M. catarrhalis* is a frequent cause of acute exacerbations of chronic obstructive pulmonary disease (Davies and Maesen, 1988; Hager *et al.*, 1987; McLeod *et al.*, 1986; Nicotra *et al.*, 1986). Invasive infections with this organism, such as bacteremia, meningitis, skeletal infections and endocarditis, are rare and occur mainly in immunocompromised individuals (Doern *et al.*, 1981; Malkamaki *et al.*, 1983).

The 1989 Consensus Report further concluded that prevention of otitis media is an important health care goal due to both its occurrence in infants and children, as well as certain populations of all age groups. In fact, the total financial burden of otitis media has been estimated to be at least 2.5 billion annually, or approximately 3% of the health care budget. Vaccines were identified as the most desired approach to the prevention of this disease for a number of reasons. For example, it was estimated that if vaccines could reduce the incidence of otitis media by 30%, this outcome could bring about an annual health care savings of at least $400 million. However, while some progress has been made in the development of vaccines for 2 of the 3 common otitis media pathogens, *Streptococcus pneumoniae and Haemophilus influenzae,* there is no indication that similar progress has been made with respect to *M. catarrhalis.* This is particularly troublesome in that *M. catarrhalis* now accounts for approximately 17-20% of all otitis media infection (Murphy, 1989).

Despite its recognized virulence potential, little is known about the mechanisms employed by *M. catarrhalis* in the production of disease or about host factors governing immunity to this pathogen. It has been documented that infections caused by *M. catarrhalis* induce an antibody response directed against this pathogen (Chapman *et al.*,

1985; Chi *et al.*, 1990; Faden *et al.*, 1994; 1992; Goldblatt *et al.*, 1990a; 1990b; Sethi *et al.*, 1995) and that some of these antibodies have the potential to exert biologic activity *via* complement-dependent killing (Chapman *et al.*, 1985) or opsonophagocytosis (Faden *et al.*, 1994). For example, an antibody response to *M catarrhalis* otitis media has been documented by means of an ELISA system using whole *M catarrhalis* cells as antigen and acute and convalescent sera or middle ear fluid as the source of antibody (Leinonen *et al.*, 1981). However, with a few exceptions (Goldblatt *et al.*, 1990b; Sethi *et al.*, 1995), the identity of the target antigen(s) remains unclear because, as in the preceding example, most studies used whole bacteria or crude outer membrane preparations as antigens to detect *M. catarrhalis*-directed antibodies.

The development of serum bactericidal antibody during *M catarrhalis* infection in adults was shown to be dependent on the classical complement pathway (Chapman *et al.*, 1985). And more recently, it was reported that young children with *M. catarrhalis* otitis media develop an antibody response in the middle ear but fail to develop systemic antibody in a uniform manner (Faden *et al.*, 1992). To date, the only *M. catarrhalis* outer membrane antigens that have been shown to be targets for antibodies that can exert protective effects in an animal model are the CopB and UspA proteins (Chen *et al.*, 1996; Helminen *et al.*, 1993a; 1994) while antibodies to protein CD have been reported to be bactericidal *in vitro* (Yang *et al.*, 1997).

The wide occurrence of *M. catarrhalis* infections and the rapid spread of β-lactamase production among clinical isolates have stimulated efforts to develop a vaccine against this pathogen (Anonymous, 1989). Several lines of evidence suggest that the induction of appropriate humoral immunity will likely be protective against respiratory tract disease caused by *M. catarrhalis.* (i) Acute *M. catarrhalis* infection induces the production of both serum and secretory antibodies against various antigenic determinants of this pathogen (Faden *et al.*, 1992; Helminen *et al.*, 1995; Sethi *et al.*, 1995). (ii) The age-dependent development of the humoral response against *M. catarrhalis* is inversely related to the prevalence of nasopharyngeal colonization and incidence of otitis media involving *M. catarrhalis* (Goldblatt *et al.*, 1990b; Vaneechoutte *et al.*, 1990). (iii) Passive immunization with *M. catarrhalis*-directed antibodies as well

as active immunization with *M. catarrhalis* outer membrane proteins enhanced pulmonary clearance of *M. catarrhalis* in an animal model (Helminen *et al.*, 1993a; 1994; Maciver *et al.*, 1993).

Previous attempts have been made to identify and characterize *M. catarrhalis* antigens that would serve as potentially important targets of the human immune response to infection (Murphy, 1989; Goldblatt *et al.*, 1990; Murphy *et al.*, 1990). Generally speaking, the surface of *M. catarrhalis* is composed of outer membrane proteins (OMPs), lipooligosaccharide (LOS) and fimbriae. *M. catarrhalis* appears to be somewhat distinct from other Gram-negative bacteria in that attempts to isolate the outer membrane of this organism using detergent fractionation of cell envelopes has generally proven to be unsuccessful in that the procedures did not yield consistent results (Murphy, 1989; Murphy and Loeb, 1989). Moreover, preparations were found to be contaminated with cytoplasmic membranes, suggesting an unusual characteristic of *the M. catarrhalis* cell envelope.

Isolation procedures have been reported for obtaining *M. catarrhalis* outer membrane components which result in what are said to be less-contaminated membrane preparations (Murphy and Loeb, 1989). Although this has allowed *M. catarrhalis* outer membrane protein profiles to be assessed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS/PAGE), these techniques did not lead to the isolation or characterization of OMPs (Murphy and Loeb, 1989). Seven or eight major OMP groups have been identified in this manner and appear to be fairly consistent between diverse *M. catarrhalis* strains. For example, OMPs have been grouped into classes A-H, beginning with bands of molecular weight around 98 kD (OMP-A) and proceeding to bands with molecular weights of about 21 kD (OMP-H) (Murphy and Loeb, 1989; Murphy, 1989).

Recently, one OMP designated CopB, was cloned and its genomic sequence and inferred amino acid sequence have been reported (U.S. Patent Application S.N. 08/025,363). *The M catarrhalis* CopB outer membrane protein appears to be fairly well conserved among strains of *M. catarrhalis,* with the CopB-specific monoclonal

antibody (MAb) 10F3 binding to the majority (70%) of 23 *M. catarrhalis* strains tested, suggesting that a strain-common epitope is present in many, but not all, strains of this pathogen. Southern blot analysis results also suggest some degree of conservation of the *copB* gene at the nucleotide sequence level among different strains. Moreover, polyclonal antisera raised against outer membrane vesicles from two MAb 10F3-unreactive strains, TTA24 and B21, reacted with the recombinant CopB protein encoded by the *copB* gene which was derived from the M. *catarrhalis* 10F3-reactive strain O35E, indicating some conservation of the protein structure of the CopB protein among the strains.

The lack of a polysaccharide capsule surrounding *M. catarrhalis* indicates that surface-exposed outer membrane antigens are the likely targets for a protective immune response. Different *M. catarrhalis* strains share remarkably similar outer membrane protein profiles (Bartos and Murphy, 1988; Murphy and Bartos, 1989), and at least three surface-exposed proteins of this organism appear to be well-conserved antigenically (Helminen *et al.*, 1993a; 1994; Hsiao *et al.*, 1995; Murphy *et al.*, 1993). One of these, the 80 kDa CopB protein (OMP B2), is a potential vaccine candidate, based on its reactivity with *M. catarrhalis* strains and enhanced pulmonary clearance of *M. catarrhalis* in a murine model (Helminen *et al.*, 1993a). It has been established that CopB expression is iron-regulated (Aebi *et al.*, 1996; Campagnari *et al.*, 1994) and that CopB is involved at some level in the ability of *M. catarrhalis* to acquire iron from human transferrin and lactoferrin (Aebi *et al.*, 1996). Expression of CopB is apparently essential for virulence of *M. catarrhalis,* at least in an animal model, because an isogenic *copB* mutant was less able than its wild-type parent strain to resist clearance from the lungs of mice (Helminen *et al.*, 1993b).

Passive immunization with polyclonal antisera raised against *M. catarrhalis* strain O35E outer membrane vesicles was also found to protect against pulmonary challenge with the heterologous *M. catarrhalis* strain TTA24. In addition, active immunization *with M. catarrhalis* outer membrane vesicles resulted in enhanced clearance of this organism from the lungs alter challenge. The positive effect of immunization in pulmonary clearance indicates that antibodies play a major role in

immunoprotection from this pathogen. In addition, the protection observed against pulmonary challenge with a heterologous *M. catarrhalis* strain demonstrates that one or more conserved surface antigens are targets for antibodies which function to enhance clearance of *M. catarrhalis* from the lungs.

With the rising importance of this pathogen in respiratory tract infections, identification of the surface components of this bacterium involved in virulence expression and immunity is becoming more important. To date, there are no vaccines available for use in humans, against any other OMP, LOS or fimbriae, that induce protective antibodies against *M. catarrhalis*. Thus, it is clear that there remains a need to identify useful antigens and which can be employed in the preparation of immunoprophylactic reagents. Additionally, once such an antigen or antigens is identified, there is a need for providing methods and compositions which will allow the preparation of these vaccines and quantities that will allow their use on a wide scale basis in prophylactic protocols.

SUMMARY OF THE INVENTION

It is a goal of the present invention to provide polypeptides and nucleic acids for CopB of *M. catarrhalis*. It is also a goal to provide methods for the diagnosis and prevention of *M. catarrhalis* infections using these polypeptide and nucleic acids, and derivatives thereof. It is a further goal of the present invention to screen for agents that may be used both in diagnosis and immunoprophylaxis of *M. catarrhalis* infections.

There is provided, in accordance with one aspect of the invention, an isolated peptide of about 5 to about 60 amino acids comprising the consecutive amino acid sequence KYAGK (SEQ ID NO:35) which is reactive with the antibody 10F3. The consecutive residues of the isolated peptide may further comprise the amino acid residues KYAGKG (SEQ ID NO:36), KYAGKGY (SEQ ID NO:37), NKYAGK (SEQ ID NO:38), NKYAGKG (SEQ ID NO:39) or NKYAGKGY (SEQ ID NO:40).

In a further embodiment, the isolated peptide may comprise at least about 20 consecutive residues of the amino acid sequence

LDIEKDKKKRTDEQLQAELDNKYAGKGY (SEQ ID NO:41) or
LDIKKDDKTLTETELQAELDNKYAGKGY (SEQ ID NO:42).

The isolated peptide may be about 5, 7, 10, 15, 20, 30, 40, 50 or 60 amino acids in length.

In another embodiment, there is provided an isolated peptide of about 20 to about 60 amino acids comprising at least about 20 consecutive residues of the amino acid sequence

LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43),
LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44) or
IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45).

Again, the isolated peptide may be about 20, 30, 40, 50 or 60 amino acids in length.

In yet another embodiment, there is provided an antigenic composition comprising (a) an isolated peptide of about 5 to about 60 amino acids comprising the amino acid sequence KYAGK (SEQ ID NO:35) and (b) a pharmaceutically acceptable buffer or diluent. Alternatively the an antigenic composition comprises (a) an isolated peptide of about 20 to about 60 amino acids comprising the amino acid sequence

LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43),
LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44) or
IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45).

and (b) a pharmaceutically acceptable buffer or diluent. The antigenic composition may further comprise a carrier conjugated to said peptide. The carrier may be KLH. The antigenic composition may further comprise an adjuvant. The antigenic composition may comprise a peptide is covalently linked to a second antigen. The second antigen may be a peptide antigen or a non-peptide antigen.

In still yet another embodiment, there is provided a vaccine composition comprising (a) an isolated peptide of about 5 to about 60 amino acids comprising the amino acid sequence KYAGK (SEQ ID NO:35) and (b) a pharmaceutically acceptable

buffer or diluent. Alternatively the an antigenic composition comprises an isolated peptide of about 20 to about 60 amino acids comprising the amino acid sequence

LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43),
LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44) or
IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45)

and a pharmaceutically acceptable buffer or diluent.

In still yet another embodiment, there is provided a method for inducing an immune response in a mammal comprising the step of providing to said mammal an antigenic composition comprising (a) an isolated peptide of about 5 to about 60 amino acids comprising the amino acid sequence KYAGK (SEQ ID NO:35) and (b) a pharmaceutically acceptable buffer or diluent. Alternatively the an antigenic composition comprises (a) an isolated peptide of about 20 to about 60 amino acids comprising the amino acid sequence

LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43),
LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44) or
IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45)

and (b) a pharmaceutically acceptable buffer or diluent.

In still yet another embodiment, there is provided an isolated and purified *M. catarrhalis* CopB antigen comprising the amino acid sequence KYAGK (SEQ ID NO:35), or in still a further embodiment

LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43),
LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44) or
IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45)

or at least about 20 consecutive residues thereof.

In still yet another embodiment, there is provided a nucleic acid encoding the CopB antigen of the M. *catarrhalis* isolates TTA24, 012E and 046E. Also provided are

copB DNA sequences of the M. *catarrhalis* isolates TTA24 (SEQ ID NO:34), 012E (SEQ ID NO:28) and 046E (SEQ ID NO:32).

In still yet another embodiment, there is provided a method for generating a strain-specific antibody to *M. catarrhalis* CopB antigen comprising the step of providing to a mammal an antigenic composition comprising (a) an isolated peptide comprising an amino acid sequence corresponding to residues of variable regions of said CopB antigen and (b) a pharmaceutically acceptable buffer or diluent.

In still yet another embodiment, there is provided a method for diagnosing *M. catarrhalis* infection comprising the step of determining the presence, in a sample, of an *M. catarrhalis,* amino acid sequence corresponding to residues of variable regions of said CopB antigen. The step of determining may comprise PCR or immunologic reactivity of an antibody with an *M. catarrhalis* antigen.

In still yet another embodiment, there is provided a method for generating a strain-common antibody to *M. catarrhalis* CopB antigen comprising the step of providing to a mammal an antigenic composition comprising (a) an isolated peptide comprising an amino acid sequence corresponding to residues of common regions of said CopB antigen and (b) a pharmaceutically acceptable buffer or diluent.

In still yet another embodiment, there is provided a method for diagnosing *M. catarrhalis* infection comprising the step of determining the presence, in a sample, of an *M. catarrhalis* amino acid sequence corresponding to residues of common regions of said CopB antigen. The step of determining may comprise PCR or comprises immunologic reactivity of an antibody with an *M. catarrhalis* antigen.

In still yet another embodiment, there is provided a method for treating an individual having an *M. catarrhalis* infection comprising providing to said individual an isolated peptide of about 20 to about 60 amino acids comprising at least the consecutive residues KYAGK (SEQ ID NO:35) or about 20 consecutive residues of the amino acid sequence

LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43),

LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44) or
IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45).

In still yet another embodiment, there is provided a method for preventing or limiting an *M. catarrhalis* infection comprising providing to a subject an antibody that reacts immunologically with an epitope formed by the amino acid sequence

KYAGK (SEQ ID NO:35),
LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43),
LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44) or
IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45).

In still yet another embodiment, there is provided a method for screening a peptide for reactivity with a CopB antibody comprising the steps of (a) providing said peptide; (b) contacting said peptide with an antibody that binds immunogically to CopB; and (c) determining the binding of said antibody to said peptide. The antibody may be 10F3, and the determining may comprise an immunoassay selected from the group consisting of a Western blot, an ELISA, an RIA and immunoaffinity separation.

In still yet another embodiment, there is provided a method for screening a peptide for the ability to induce a protective immune response against *M. catarrhalis* comprising the steps of (a) providing said peptide; (b) administering a peptide in a suitable form to an experimental animal; (c) challenging said animal with *M. catarrhalis;* and (d) assaying the infection of said animal with *M. catarrhalis.* The animal may be a mouse, the challenge is a pulmonary challenge, and the assaying comprises assessing the degree of pulmonary clearance by said mouse. Preferably, the peptide is a CopB peptide, and more preferably, the CopB peptide encompasses residues 296-300 of *M. catarrhalis* strain 035E. Alternatively, the preferred embodiment may employ a CopB peptide is a peptide having at least 6 consecutive amino acids from *M. catarrhalis* strain 035E, including residue 295.

In still yet another embodiment, there is provided an isolated peptide having at least about 5 consecutive amino acids from the CopB protein of *M. catarrhalis* and

including residues 296-300, or the analogous position thereof when compared to strain 035E or having about 6 consecutive amino acids from the CopB protein of *M. catarrhalis,* wherein said peptide includes residue 295, or the analogous position thereof when compared to strain 035E. The isolated peptide may be between 5 and 60 amino acids in length. The peptide may comprise non-CopB sequences and even non-*M. catarrhalis* sequences.

In still yet another embodiment, there is provided an antigen composition comprising (i) an isolated peptide having at least about 6 consecutive amino acids from the CopB protein of *M. catarrhalis,* wherein said peptide includes residue 295, or the analogous position thereof when compared to strain 035E and (ii) a pharmaceutically acceptable buffer or diluent.

BRIEF DESCRIPTION OF THE DRAWINGS

The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present invention. The invention may be better understood by reference to one or more of these drawings in combination with the detailed description of specific embodiments presented herein.

FIG. 1A and FIG. 1B Western blot of oligopeptide (region 1) antisera against CopB antigen from each of strains 035E, 01 2E, 046E and TTA24. Outer membrane vesicles of the *M. catarrhalis* strains 035E (lane 1), 012E (lane 2), TTA24 (lane 3) and 046E (lane 4) probed with MAb 10F3 (FIG. 1A) or mouse antiserum raised against peptide from region 1 conjugated with KLH (FIG. 1B).

FIG. 2 Amino acid sequence and reactivity with antibodies of both region 1 from CopB of four different *M. catarrhalis* strains and synthetic peptides derived from these sequences. The numbers (*i.e.*, 275-300) above the region 1 sequences indicate the amino acid positions in the intact CopB protein. The numbers (*i.e.*, 1-26) above the R1 synthetic peptide indicate residue positions in the R1 peptide and its derivatives. Boxes indicates residues differing from those in 035E (SEQ ID NO:1). The reactivity of MAb 10F3, GST-26 antiserum, and GST antiserum with the CopB proteins of these strains in

western blot analysis is indicated on the right; these two antisera were diluted 1:100 for use in this assay. The reactivity of the synthetic peptides with these same antibody preparations in the ELISA is indicated on the right opposite these peptides. The asterisk indicates that the ELISA with MAb 10F3 was performed after adsorption of the peptides to the microtiter wells in the presence of glutaraldehyde. The ELISA readings (OD_{410}) were recorded after 30 min as described in the Materials and Methods of Example V.

FIG. 3 Fine mapping of the MAb 10F3-reactive epitope using a dot blot assay involving immobilized synthetic decapeptides. Binding of MAb 10F3 to individual peptides was detected by incubating the membrane with radioiodinated goat anti-mouse immunoglobulin. The autoradiograph of the dot blot (peptides 1-8) is shown on the left, the corresponding amino acid sequence of each of the overlapping decapeptides with their respective positions in the CopB protein is listed in Table V. The numbers written vertically denote the amino acid positions in the intact CopB protein; the R1 region is underlined.

FIG. 4 Hydropathy plot and secondary structure analysis of region 1 from the CopB protein of *M. catarrhalis* strain 035E. The hydropathy plot was generated by the method of Kyte and Doolittle (1982) using a window size of seven residues and secondary structure predictions were made using the method of Chou and Fasman (1978) as found in MacVector® sequence analysis software (Version 6.0). The numbers on the horizontal scale denote the amino acid positions in the intact CopB protein; the R1 region is underlined.

FIG. 5A and FIG. 5B Western blot analysis of the reactivity of MAbs and polyclonal serum antibodies with outer membrane proteins of *M. catarrhalis* strains, the GST-26 fusion protein, and GST alone. Proteins present in outer membrane vesicles from strains 035E (lane 1), 012E (lane 2), TTA24 (lane 3), and 046E (lane 4), purified GST-26 (lane 5) and GST (lane 6) were resolved by SDS-PAGE, transferred to nitrocellulose, and probed with MAb 10F3 (FIG. 5A) or with mouse GST-26 antiserum (dilution 1:1000) (FIG. 5B). The CopB protein from strain 035E gave rise to a doublet

reactive with MAb 10F3 in this study. Molecular weight position markers are shown on the left sides of the blots.

FIGS. 6A-6D Bactericidal activity of MAb 10F3 against *M. catarrhalis* strains. Suspensions of the MAb 10F3-reactive strains 035E (FIG. 6A) and 012E (FIG. 6B) and the MAb 10F3-unreactive strains TTA24 (FIG. 6C) and 046E (FIG. 6D) were incubated with: normal human serum (closed triangles), normal human serum and MAb 10F3 (open squares), heat-inactivated normal human serum and MAb 10F3 (closed squares), normal human serum and GST-26 antiserum (open circles) or normal human serum and GST antiserum (closed circles). Portions of the reaction mixture were removed over time and spread on BHI agar plates to determine the number of viable bacteria. Note that the open circles and open squares in FIG. 6B are superimposed.

DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

The present invention relates to the identification of CopB outer membrane protein sequences from multiple strains of *M. catarrhalis,* and the nucleic acid sequences corresponding thereto. In addition, the present invention provides insights into the antigenic structure of CopB, based on an analysis of the amino acid sequences. This information is useful in developing both strain-specific and strain-common reagents for the diagnosis of *M. catarrhalis* infection. In addition, an understanding of which regions of the molecule are conserved permits the development of agents that will be useful in protecting against *M. catarrhalis* infections, *e.g.,* in the preparation of vaccine reagents. Particular embodiments relate to peptides, polypeptides and nucleic acids corresponding to the CopB molecules, peptides and antigenic compositions derived therefrom, and methods for the diagnosis and treatment of *M. catarrhalis* disease.

As stated above, *M. catarrhalis* infections present a serious health challenge, especially to the young. CopB represents an important antigenic determinant, as the monoclonal antibody (MAb) termed 10F3 has been shown to protect experimental animals, as measured in a pulmonary clearance model, when provided in passive immunizations. The present invention, by virtue of new information regarding the structure of CopB provides such improved compositions and methods.

In a first embodiment, the present invention provides for the identification of conserved and *non-conserved* portions of the CopB polypeptide. A homology comparison illustrates at least ten discrete regions of low homology. These regions are designated region 1 (about residues 275-302), region 2 (about residues 322-329), region 3 (about residue 348), region 4 (about residues 400-416), region 5 (about residues 437-450), region 6 (about residues 485-486), region 7 (about residues 526-541), region 8 (about residues 583-596), region 9 (about residues 637-640) and region 10 (about residues 683-695). The numbering is based on the amino acid sequence of 035E strain (SEQ ID NO:29) which is encoding by SEQ ID NO:30 beginning at position 14.

The remaining regions of the molecule may be considered "conserved." Of particular note in this regard are the regions from about residue 10 to about residue 270, from about residue 350 to about residue 395, from about residue 455 to about residue 480, from about residue 490 to about residue 520, from about residue 545 to about residue 580, from about residue 600 to about residue 635 and from about residue 645 to about residue 680. Again, the numbering is based on the 035E strain (SEQ ID NO:29).

In a second embodiment, the present invention has identified the specific epitope to which 10F3 binds. A peptide sequence, found between residues 275-302 (region 1; SEQ ID NO:1) of the CopB protein of *M. catarrhalis* strain 035E, is sufficient to be bound by 10F3. In addition, this same region from the CopB protein from *M. catarrhalis* strain 012E (SEQ ID NO:2) is bound by 10F3, although there are differences in the identified region between these two strains. Two other strains, 046E and TTA24, are not bound by 10F3. The sequences from the identified region of these strains, SEQ ID NO:4 from strain TTA24 and SEQ ID NO:5 from strain O35E, respectively, contain differences in common to each other but not in common with 035E or 012E. It is clear that this region plays an important role in the biology of the pathogen and, from this information, one can deduce amino acids residues that are critical to and not critical to 10F3 antibody binding.

In a third embodiment of the invention, the specific epitope to which 10F3 binds is even further defined. Peptides LDNKYAGKGY (SEQ ID NO:25; peptide 7; residues

293-302 of the CopB protein of strain O35E) and KYAGKGYKLG (SEQ ID NO:26; peptide 8; residues 296-305 of the CopB protein of strain O35E) specifically bind 10F3; whereas, peptides 1 - 6 (FIG. 3; Table V) exhibit non-specific binding in the presence of 10F3 and do not appear to have any specific affinity for 10F3. Thus, the small peptide KYAGK (SEQ ID NO:35) is preferred for binding to 10F3, a more preferred peptide is KYAGKG (SEQ ID NO:36) and an even more preferred peptide is KYAGKGY (SEQ ID NO:37).

In a fourth embodiment of the invention, the peptide sequence, LDNKYAGKGY (SEQ ID NO:25) allows maximal binding of MAb 10F3 compared to other peptides examined herein. The asparagine (N) residue (position 3 of SEQ ID NO:25) appears to improve binding. Although it is recognized that other residues may be inserted at this position and also provide similar or even greater binding to 10F3, it is also clear that this position influences the binding kinetics of the epitope, because when aspartic acid (D) is substituted for N at this position, reactivity with 10F3 is abolished. Thus, based on the data presented in FIG. 2 and FIG. 3, another preferred small peptide for binding to 10F3 is NKYAGK (SEQ ID NO:38), a more preferred peptide is NKYAGKG (SEQ ID NO:39) and an even more preferred peptide is NKYAGKGY (SEQ ID NO:40). Further embodiments of the present invention are discussed below.

Several lines of evidence indicate that portions of the CopB protein are exposed on the surface of *M. catarrhalis*. These include (i) the reactivity of the CopB-specific MAb 10F3 with a surface-exposed epitope (Helminen *et al.*, 1993a), (ii) the bactericidal activity of MAb 10F3 against strains that react with this MAb (FIGS. 6A-6D), and (iii) the likely function of the CopB protein in a TonB-dependent physiologic process (Aebi *et al.*, 1996). In addition, FrpB, an iron-regulated outer membrane protein of *Neisseria meningitidis* that is 49% identical to the CopB protein of strain 035E, has also been shown to be a target for bactericidal antibody (Pettersson *et al.*, 1990; 1995).

Comparison of the deduced amino acid sequence of the CopB proteins from the four *M. catarrhalis* strains revealed that reactivity with MAb 10F3 was not an indicator of similarity between CopB proteins. Furthermore, the primary amino acid sequence of

this protein was very well conserved among these four strains. The level of identity among these four CopB proteins proved to be greater than 90%, a finding which is somewhat encouraging from the standpoint of vaccine development. A similar level of amino acid sequence conservation has been reported for the CD outer membrane protein of *M. catarrhalis* (Hsiao *et al.*, 1995).

The facts that the MAb 10F3-reactive epitope was detected in the majority of *M. catarrhalis* strains (Helminen *et al.*, 1993a) and that this epitope induced the synthesis of antibodies that were biologically active both *in vivo* (Helminen *et al.*, 1993a) and *in vitro* (FIGS. 6A-6D) made the determination of this epitope's location within the CopB protein an important step in the evaluation of CopB as a potential vaccine candidate. This task was facilitated considerably by the finding that the deduced amino acid sequence differences between the CopB protein of the MAb 10F3-reactive strain 035E and that of the MAb 10F3-unreactive strain TTA24 were limited to five well-defined regions, one of which had to contain the MAb 10F3-reactive epitope. The analyses of the *copB* genes from two other *M. catarrhalis* strains reinforced the likelihood that one of these five regions bound MAb 10F3 and prompted the synthesis of oligopeptides that contained the corresponding amino acid sequences from the CopB protein of strain 035E.

The use of synthetic peptides in an ELISA clearly identified region 1 as the sequence containing the MAb 10F3-reactive epitope. Unexpectedly, binding of MAb 10F3 to peptide R1 (SEQ ID NO:1; FIG. 2) required the presence of the cross-linking agent glutaraldehyde during coating of this peptide to the polystyrene surface of the ELISA plate. At first, it was thought that either inherently poor binding of R1 (SEQ ID NO:1) to the microtiter well or binding of R1 (SEQ ID NO:1) to the well *via* the MAb 10F3-reactive residues was responsible for this finding. However, the failure of soluble R1 (SEQ ID NO:1) to bind MAb 10F3 in an inhibition assay as well as the subsequent observation that GST-26 antiserum readily bound to microtiter wells coated with R1 (SEQ ID NO:1) in the absence of glutaraldehyde invalidated these hypotheses and made conformational changes of R1 (SEQ ID NO:1) caused by glutaraldehyde a more plausible explanation. Glutaraldehyde preferentially cross-links amino groups of lysine residues such as those present in both the N- and C-terminal portions of R1 (SEQ ID

NO:1; FIG. 2). Crosslinking of these lysine residues may have led to the formation of a loop-like structure that more closely resembled the native conformation of the MAb 10F3-reactive epitope. This hypothesis is supported by the lack of binding of MAb 10F3 by peptide R1E (SEQ ID NO:6; FIG. 2) which, as was subsequently determined, contained the binding site for MAb 10F3 but lacked the lysine-rich N-terminal portion of R1.

Differential binding of MAb 10F3 to the peptides R1A (SEQ ID NO:7; FIG. 2) and R1B (SEQ ID NO:3; FIG. 2) indicated that this MAb likely bound the C-terminal portion of R1 (SEQ ID NO:1) and that the asparagine (N) residue in position 21 was essential for binding. This assumption was confirmed by fine mapping of the MAb 10F3-reactive epitope using overlapping decapeptides (FIG. 3) which demonstrated optimal binding of MAb 10F3 to a decapeptide (peptide 7; SEQ ID NO:25) that corresponded to the residues 19-28 of region 1 (*i.e.*, residues 293-302 of the CopB protein). This epitope comprises the C-terminal portion of a highly hydrophilic domain when analyzed by means of the hydropathy algorithm of Kyte and Doolittle (1982) (FIG. 4). This finding is in keeping with the proven, bacterial cell surface-exposed location of the binding site for MAb 10F3. It is also noteworthy that the MAb 10F3-reactive epitope appears to be part of the largest variable region of the otherwise highly conserved CopB proteins studied here. This finding is reminiscent of other well-studied bacterial outer membrane proteins, such as the major outer membrane protein P2 of nontypeable *Haemophilus influenzae,* for which it has been demonstrated that some surface-exposed domains exhibit strain-specific antigenicity (Duim *et al.*, 1994; 1996).

Characterization of the immune response directed against the GST-26 fusion protein containing the MAb 10F3 binding site revealed that, in western blot analysis, antibodies in the GST-26 antiserum bound the CopB protein of strain 035E, but not the CopB protein of the equally MAb 10F3-reactive strain 012E (FIG. 5B). This result suggested that these polyclonal antibodies did not possess the same antigenic specificity as MAb 10F3. In fact, the strong reactivity of the GST-26 antiserum with the MAb 10F3-unreactive peptides R1B (SEQ ID NO:3; FIG. 2) and R1-TTA24 (SEQ ID NO:4; FIG. 2) suggested that most of these polyclonal antibodies are directed against epitopes

of peptide R1 (SEQ ID NO:1) other than that which binds MAb 10F3. The reduced reactivity of GST-26 antiserum with peptide R1E (SEQ ID NO:6), which consisted of the 15 C-terminal residues of peptide R1, indicated that, not unexpectedly, some of the antibodies in the GST-26 antiserum are directed against epitopes located in the N-terminal portion of peptide R1.

Antibodies raised against GST-26 bound to the surface of whole cells of *M. catarrhalis* strains 035E, 012E and TTA24 (Table VI), but exerted bactericidal activity against only strain 012E (FIGS. 6A-6D). That these polyclonal antibodies, raised against this strain 035E-derived region 1 sequence (FIG. 2) could bind to strains 012E and TTA24 is not surprising when one considers that this same region of the CopB proteins of these other two strains was about 73%-80% identical to that of strain 035E (FIG. 2). It is also likely that these polyclonal antibodies had relatively low binding affinities which precluded their detection in western blot analysis. However, the fact that these same surface-directed antibodies killed only strain 012E and not the homologous strain 035E or the heterologous strain TTA24 is more difficult to explain. One possibility is the presence of an immunodominant determinant in the R1 region that induced the synthesis of blocking antibodies that bound to the R1 region of CopB of strains 035E and TTA24 but which were unreactive with the R1 region of CopB of strain 012E. Alternatively, there may have been other blocking antibodies in the mouse antiserum or in the normal human serum used as the source of complement in these studies; these blocking antibodies could be directed against a surface determinant of strains 035E and TTA24 that was absent on strain 012E.

In summary, comparison of the deduced amino acid sequences of the CopB proteins from four strains of *M. catarrhalis* revealed a striking degree of identity among these macromolecules. This conservation of the primary amino acid sequence of the CopB protein among strains of this pathogen was crucial for localization of the protective epitope on the CopB protein of *M. catarrhalis* strain 035E that is recognized by the bactericidal MAb 10F3. Immunization with a fusion protein reactive with MAb 10F3 induced the synthesis of CopB-specific antibodies bactericidal for at least one MAb 10F3-reactive strain of *M. catarrhalis*.

It is important to note that screening methods for diagnosis and prophylaxis are readily available, as set forth below. Thus, the ability to (i) test peptides, mutant peptides and antibodies for their reactivity with each other and (ii) test peptides and antibodies for the ability to prevent infections *in vivo,* provide powerful tools to develop clinically important reagents.

1. CopB Peptides and Polypeptides

The present invention, in one embodiment, encompasses the three new amino acid sequences presented in SEQ ID NO:27, SEQ ID NO:31 and SEQ ID NO:33. Also encompassed in the present invention are hybrid molecules containing portions from the CopB of one strain fused with portions of the CopB from another strain. For example, one may fuse residues 1-350 of strain TTA24 CopB with residues 351-759 of strain 035E CopB. Alternatively, a fusion could be generated with sequences from three or even four strains represented in a single CopB molecule, thus combining epitopes from each of the various strains. Also encompassed are fragments of the disclosed CopB molecules, as well as insertion, deletion or replacement mutants in which non-CopB sequences are introduced, CopB sequences are removed, or CopB sequences are replaced by or fused with non-CopB sequences, respectively. For example, in order to improve expression in certain systems, it is necessary to "fuse" amino-terminal regulatory sequences to the peptide of interest.

CopB, according to the present invention, may be advantageously cleaved into fragments for use in further structural or functional analysis, or in the generation of reagents such as CopB-related polypeptides and CopB-specific antibodies. This can be accomplished by treating purified or unpurified CopB with a peptidase such as endoproteinase glu-C (Boehringer, Indianapolis, IN). Treatment with CNBr is another method by which CopB fragments may be produced from natural CopB. Recombinant techniques also can be used to produce specific fragments of CopB.

More subtle modifications and changes may be made in the structure of the encoded CopB polypeptides of the present invention while retaining a molecule that encodes a protein or peptide with characteristics of natural or "wild-type" CopB

polypeptide. The following is a discussion based upon changing of the amino acids of a CopB protein to create an equivalent, or even an "improved," second-generation molecule. The amino acid changes may be achieved by changing the codons of the DNA sequence, according to the following codon table:

Table I

Amino Acid Names and abbreviations			Codons					
Alanine	Ala	A	GCA	GCC	GCG	GCU		
Cysteine	Cys	C	UGC	UGU				
Aspartic acid	Asp	D	GAC	GAU				
Glutamic acid	Glu	E	GAA	GAG				
Phenylalanine	Phe	F	UUC	UUU				
Glycine	Gly	G	GGA	GGC	GGG	GGU		
Histidine	His	H	CAC	CAU				
Isoleucine	Ile	I	AUA	AUC	AUU			
Lysine	Lys	K	AAA	AAG				
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
Methionine	Met	M	AUG					
Asparagine	Asn	N	AAC	AAU				
Proline	Pro	P	CCA	CCC	CCG	CCU		
Glutamine	Gln	Q	CAA	CAG				
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Threonin e	Thr	T	ACA	ACC	ACG	ACU		
Valine	Val	V	GUA	GUC	GUG	GUU		
Tryptophan	Trp	W	UGG					
Tyrosine	Tyr	Y	UAC	UAU				

Generally, it is known that certain amino acids may be substituted for other amino acids in a protein structure in order to modify or improve its antigenic or immunogenic activity (see, *e.g.*, Kyte & Doolittle, 1982; Hopp, U.S. Patent 4,554,101, incorporated herein by reference). For example, through the substitution of alternative amino acids, small conformational changes may be conferred upon a polypeptide which result in increased activity or stability. Alternatively, amino acid substitutions in certain polypeptides may be utilized to provide residues which may then be linked to other molecules to provide peptide-molecule conjugates which retain enough antigenicity of

the starting peptide to be useful for other purposes. For example, a selected CopB peptide that may be covalently bound to a solid support may be constructed which would have particular advantages in diagnostic embodiments.

The importance of the hydropathic index of amino acids in conferring interactive biological function on a protein has also been discussed generally by Kyte & Doolittle (1982), wherein it is found that certain amino acids may be substituted for other amino acids having a similar hydropathic index or core and still retain a similar biological activity. As displayed in Table II below, amino acids are assigned a hydropathic index on the basis of their hydrophobicity and charge characteristics. It is believed that the relative hydropathic character of the amino acid determines the secondary structure of the resultant protein, which in turn defines the interaction of the protein with substrate molecules. Preferred substitutions which result in an antigenically equivalent CopB peptide or CopB protein will generally involve amino acids having index scores within ±2 units of one another, and more preferably within ±1 unit, and even more preferably, within ±0.5 units. Thus, for example, isoleucine, which has a hydropathic index of ±4.5, will preferably be exchanged with an amino acid such as valine (+4.2) or leucine (+3.8). Alternatively, at the other end of the scale, lysine (-3.9) will preferably be substituted for arginine (-4.5), and so on.

TABLE II

Amino Acid	Hydropathic Index
Isoleucine	4.5
Valine	4.2
Leucine	3.8
Phenylalanine	2.8
Cysteine/cystine	2.5
Methionine	1.9
Alanine	1.8
Glycine	-0.4
Threonine	-0.7
Tryptophan	-0.9
Serine	-0.8
Tyrosine	-1.3
Proline	-1.6
Histidine	-3.2
Glutamic Acid	-3.5
Glutamine	-3.5
Aspartic Acid	-3.5
Asparagine	-3.5
Lysine	-3.9
Arginine	-4.5

Substitution of like amino acids may also be made on the basis of hydrophilicity, particularly where the biological functional equivalent protein or peptide thereby created is intended for use in immunological embodiments. U.S. Patent 4,554,101, incorporated herein by reference, states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with its immunogenicity and antigenicity, *i.e.* with an important ,biological property of the protein.

As detailed in U.S. Patent 4,554,101, each amino acid has also been assigned a hydrophilicity value. These values are detailed below in Table III.

TABLE III

Amino Acid	Hydropathic Index
arginine	+3.0
lysine	+3.0
aspartate	+3.0 ± 1
glutamate	+3.0 ± 1
serine	+0.3
asparagine	+0.2
glutamine	+0.2
glycine	0
threonine	-0.4
alanine	-0.5
histidine	-0.5
proline	0.5 ± 1
cysteine	-1.0
methionine	-1.3
valine	-1.5
leucine	-1.8
isoleucine	-1.8
tyrosine	-2.3
phenylalanine	2.5
tryptophan	-3.4

It is understood that one amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent CopB peptide or protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ±2 is preferred, those

which are within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred.

Accordingly, these amino acid substitutions are generally based on the relative similarity of R-group substituents, for example, in terms of size, electrophilic character, charge, and the like. In general, preferred substitutions which take various of the foregoing characteristics into consideration will be known to those of skill in the art and include, for example, the following combinations: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

In addition, peptides derived from these polypeptides, including peptides of at least about 6 consecutive amino acids from these sequences, are contemplated. Alternatively, such peptides may comprise about 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59 or 60 consecutive residues. For example, a peptide that comprises 6 consecutive amino acid residues may comprise residues 1 to 6, 2 to 7, 3 to 8 and so on of the CopB protein. Such peptides may be represented by the formula

x to (x + n) = amino to carboxy residues of the first and last consecutive residues

where x is equal to any number from 1 to the full length of the CopB protein minus n, and n is equal to the length of the peptide minus 1. So, for CopB of TTA24, x = 1 to (759 -n), for CopB of 012E, x = 1 to (758 - n), for CopB of 035E, x = 1 to (759 - n) and for CopB of 046E, x = 1 to (767 - n). Where the peptide is 10 residues long (n = 10-1), the formula represents every 10-mer possible for each antigen., For example, where x is equal to 1 the peptide would comprise residues 1 to (1 + [10-1]), or 1 to 10. Where x is equal to 2, the peptide would comprise residues 2 to (2 + [10-2]), or 2 to 11, and so on, until the peptide would comprise residues 750 to (750 + [759 - 750]), or 750 to 759.

Syntheses of CopB peptides are readily achieved using conventional synthetic techniques such as the solid phase method *(e.g.,* through the use of a commercially

available peptide synthesizer such as an Applied Biosystems Model 430A Peptide Synthesizer). Peptides synthesized in this manner may then be aliquoted in predetermined amounts and stored in conventional manners, such as in aqueous solutions or, even more preferably, in a powder or lyophilized state pending use.

In general, due to the relative stability of peptides, they may be readily stored in aqueous solutions for fairly long periods of time if desired, *e.g.*, up to six months or more, in virtually any aqueous solution without appreciable degradation or loss of antigenic activity. However, where extended aqueous storage is contemplated it will generally be desirable to include agents including buffers such as Tris or phosphate buffers to maintain a pH of 7.0 to 7.5. Moreover, it may be desirable to include agents which will inhibit microbial growth, such as sodium azide or Merthiolate. For extended storage in an aqueous state it will be desirable to store the solutions at 4°C, or more preferably, frozen. Of course, where the peptide(s) are stored in a lyophilized or powdered state, they may be stored virtually indefinitely, *e.g.*, in metered aliquots that may be rehydrated with a predetermined amount of water (preferably distilled, deionized) or buffer prior to use.

Of particular interest are peptides that represent epitopes that lie within the CopB antigens of the present invention. An "epitope" is a region of a molecule that stimulates a response from a T-cell or B-cell, and hence, elicits an immune response from these cells. An epitopic core sequence, as used herein, is a relatively short stretch of amino acids that is structurally and biochemically "complementary" to, and therefore will bind to, binding sites on antibodies or T-cell receptors.

Antibody binding sites also may be termed "linear" epitopes, as the residues that are involved in the actual physical interaction of antigen and antibody are linear, or at least in a discrete region of the molecule. Other epitopes may be non-linear in that widely separated regions of the molecule participate in the interaction. Operationally, these two types of epitopes may be distinguished on the basis of binding in denaturing Western blot analysis - positive binding indicates linear, negative binding indicates non-linear.

The identification of epitopic core sequences is known to those of skill in the art. For example U.S. Patent 4,554,101 teaches identification and preparation of epitopes from amino acid sequences on the basis of hydrophilicity, and by Chou-Fasman analyses. Numerous computer programs are available for use in predicting antigenic portions of proteins, examples of which include those programs based upon Jameson-Wolf analyses (Jameson and Wolf, 1988; Wolf *et al.*, 1988), the program PepPlot® (Brutlag *et al.*, 1990; Weinberger *et al.*, 1985), and other new programs for protein tertiary structure prediction (Fetrow & Bryant, 1993) that can be used in conjunction with computerized peptide sequence analysis programs. Epitopic core sequences of the present invention also may be operationally defined in terms of their ability to compete with or perhaps displace the binding of the corresponding CopB antigen to the corresponding CopB-directed antisera.

For any of the variant molecules described above, immunoassays may be employed to screen for reactivity of these molecules with previously developed immunoreagents. Similarly, new antibodies and antisera, based on these variants, may be generated in order test their reactivity with existing "native" or "wild-type" molecules. Moreover, the availability of an art-recognized model for M. *catarrhalis* provides a ready means for testing variant *in vivo* for their ability to protect subjects from M. *catarrhalis* infections.

2. A Structural Analysis of the 10F3 Epitope of CopB

Based on the foregoing discussion, it is clear that it should be possible to substitute specific amino acids of the 10F3 epitope of region 1 with other amino acids and still retain, and possibly even improve upon, the reactivity of the epitope with 10F3 or other antibodies that recognize the CopB protein of *Moraxella*. Which amino acids may be interchanged is dependent upon several factors, including but not limited to, the hydropathicity, hydrophilicity, polarity and tertiary structure of the targeted amino acid as well as other amino acids proximal to the targeted residue which may influence its characteristics. In order to illustrate how one would begin to determine which amino acids are likely candidates for substitution in order to create new peptides, polypeptides

or proteins that are either strain-specific or strain-common antigens for *Moraxella*, the analysis of variable region 1 is described below.

It appears that the fourth position of region 1 may be at least either glutamate or lysine and not adversely affect reactivity with 10F3. Both of these amino acids are charged residues, but glutamate is negative and lysine is positive; therefore, it is reasonable to speculate that any charged residue may be present in this position and still result in a protein or peptide that binds to the 10F3 antibody. This data suggests that the residue at position 4 is unlikely to be critical for strain-specific antibody binding. But as indicated above, other factors may affect the ability of one amino acid to be substituted for another without causing an undesired effect. Mutated peptides, polypeptides and proteins resulting from such substitutions could be easily screened for reactivity using the immunoassays described in the instant specification.

In strains 035E and 012E that bind 10F3, position 6 is occupied by aspartate, a negatively charged amino acid. In strains TTA24 and 046E which do not bind 10F3, the homologous position is occupied by either a positively charged lysine or a weakly basic asparagine. If one were to only consider the charge of the amino acids with regard to antibody binding then one would deduce that position 6 may be critical to reactivity with 10F3; however, data disclosed herein show that position 6 alone does not determine reactivity with 10F3. Other factors appear to be more important for 10F3 binding. But amino acids at position 6 could have a synergistic effect upon reactivity with 10F3 or other antibodies. Such an assertion could be tested as indicated.

Position 7 may be occupied by either lysine or aspartate and still result in a peptide that binds to 10F3. Given the analysis of positively and negatively charged residues described previously, it is reasonable to deduce that any one of lysine, arginine, histidine, aspartate or glutamate could also be present at the seventh position of a 10F3-binding peptide. Further residues may also occupy this position and still result in a peptide that retains 10F3 antibody binding.

Position 9 is occupied by either a positively charged lysine or a neutral amino acid in both strains that react or do not react with 10F3. It is possible that other

positively charged or even neutral amino acids may occupy position 9. At position 10 of each of the currently known 10F3-binding species there is a positively charged amino acid. However, as 035E contains arginine at this position and 012E contains lysine, it seems reasonable to propose that histidine, another positively charged amino acid, could also be present at this position in a 10F3-binding peptide.

Similarly, at position 12 there are one of two negatively charged amino acids. Therefore, interchange of aspartate or glutamate at this position is contemplated to be acceptable. Furthermore, it is not unreasonable to at least propose that a positively charged amino acid may be present at this position. Although not representing the type of change that is already supported from data generated from a known CopB strain, the effect of inserting a positively charged residue at this position can be easily addressed. In these circumstances all that would be required would be to simply insert a positively charged amino acid at this position and to test the resultant peptide in any one of a number of acceptable immunobinding assays and to determine whether the resultant peptide retained the capability to bind to the MAb 10F3.

Variability also occurs around position 13 of the region where a negatively charged glutamate or neutral threonine is present and 10F3 binding reactivity is still retained. As neutral amino acids are also present in TTA24 and 046E it is likely that position 13 does not play a critical role in determining reactivity with 10F3, and any one of several neutral amino acids, including the aromatic phenylalanine and tryptophan, or negatively charged residues may occupy position 13 without causing undesirable effects. The same may be true of positions 14, 16 and 24 which are also occupied by either negatively charged or neutral residues in all four strains.

Position 20 is occupied by either the negatively charged aspartate or by the weakly basic asparagine in strains that bind 10F3, but as this position is also occupied by aspartate in the strains that do not bind 10F3 little can be deduced about its importance in the reactivity of the region. But as indicated earlier, its role could be easily determined by preparing mutant peptides, polypeptides or proteins and screening them for reactivity with 10F3 or another antibody as described in the instant specification.

Position 21 is occupied by only the weakly basic asparagine in the 035E and 012E strains. In TTA24 and 046E it is occupied by the negatively charged aspartate and the positively charged lysine, respectively. Although there are significant differences in charge between the residues, there is no pattern other than charge. Nonetheless, data described herein indicate that position 21 does indeed contain the key residue for determining maximal binding to 10F3.

Positions 22 and 26 of the region are both occupied by either the positively charged lysine or by one of two negatively charged residues. As there is no apparent relationship to 10F3 binding associated with either positively or negatively charged residues at this position. It is. reasonable to assume that these residue do not play a important roles in strain-specific reactivity given their current, surrounding residues. But if proximal residues are altered, then residues at either of these positions may exert an effect on antibody binding.

3. CopB Nucleic Acids

In addition to polypeptides, the present invention also encompasses nucleic acids encoding the CopB antigens of SEQ ID NO:27, SEQ ID NO:31 and SEQ ID NO:33 which are provided as SEQ ID NO:28, SEQ ID NO:32 and SEQ ID NO:34, respectfully. Because of the degeneracy of the genetic code, many other nucleic acids also may encode a given CopB, however. For example, four different three-base codons encode the amino acids alanine, glycine, proline, threonine and valine, while six different codons encode arginine, leucine and serine. Only methionine and tryptophan are encoded by a single codon. Table I provides a list of amino acids and their corresponding codons for use in such embodiments. In order to generate any nucleic acid encoding a given CopB, one need only refer to the codon table provided herein. Substitution of the natural codon with any codon encoding the same amino acid will result in a distinct nucleic acid that encodes CopB. As a practical matter, this can be accomplished by site-directed mutagenesis of an existing *copB* gene or *de novo* chemical synthesis of one or more nucleic acids.

These observations regarding codon selection, site-directed mutagenesis and chemical synthesis apply with equal force to the discussion of substitutional mutants CopB peptides and polypeptides, as set forth above. More specifically, substitutional mutants generated by site-directed changes in the nucleic acid sequence that are designed to alter one or more codons of a given polypeptide or epitope may provide a more convenient way of generating large numbers of mutants in a rapid fashion.

The nucleic acids of the present invention provide for a simple way to generate fragments *(e.g.,* trunications) of CopB, CopB-CopB fusion molecules (discussed above) and CopB fusions with other molecules. For example, utilization of restriction enzymes and nuclease in the *copB* gene permits one to manipulate the structure of these genes, and the resulting gene products. Vectors and cloning strategies are well known to those of skill in the art.

The nucleic acid sequence information provided by the present disclosure also allows for the preparation of relatively short DNA (or RNA) sequences that have the ability to specifically hybridize to gene sequences of the selected *copB* gene. In these aspects, nucleic acid probes of an appropriate length are prepared based on a consideration of the coding sequence of the *copB* gene, or flanking regions near the *cop B* gene, such as regions downstream and upstream in the *M. catarrhalis* chromosome. The ability of such nucleic acid probes to specifically hybridize to *copB* gene sequences lends them particular utility in a variety of embodiments. For example, the probes can be used in a variety of diagnostic assays for detecting the presence of pathogenic organisms in a given sample. In addition, these oligonucleotides can be inserted, in frame, into expression constructs for the purpose of screening the corresponding peptides for reactivity with existing antibodies or for the ability to generate diagnostic or prophylactic reagents. Finally, as discussed above, they may be utilized to generate mutants by site-directed mutagenesis.

To provide certain of the advantages in accordance with the invention, the preferred nucleic acid sequence employed for hybridization studies or assays includes sequences that are complementary to at least a 10 to 20, or so, nucleotide stretch of the

sequence, although sequences of 30 to 60 or so nucleotides are also envisioned to be useful. Such molecules may be term oligonucleotides or probes. A size of at least 9 nucleotides in length helps to ensure that the fragment will be of sufficient length to form a duplex molecule that is both stable and selective. Molecules having complementary sequences over stretches greater than 10 bases in length are generally preferred, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of the specific hybrid molecules obtained. Thus, one will generally prefer to design nucleic acid molecules having *copB* gene-complementary stretches of 15 to 20 nucleotides, or even longer, such as 30 to 60, where desired. Such fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, by application of nucleic acid reproduction technology, such as the PCR technology of U.S. Patent 4,603,102, or by introducing selected sequences into recombinant vectors for recombinant production.

The oligos or probes that would be useful may be derived from any portion of the sequences of strains TTA24 (SEQ ID NO:34), 012E (SEQ ID NO:28) or 046E (SEQ ID NO:32). Therefore, oligos or probes are specifically contemplated that comprise nucleotides 1 to 9, or 2 to 10, or 3 to 11 and so forth up to a probe comprising the last 9 nucleotides of the nucleotide sequence of TTA24 (SEQ ID NO:34), 012E (SEQ ID NO:28) or 046E (SEQ ID NO:32). Thus, each probe would comprise at least about 9 linear nucleotides of the nucleotide sequence of TTA24 (SEQ ID NO:34), 012E (SEQ ID NO:28) or 046E (SEQ ID NO:32), designated by the formula "n to n + 8," where n is an integer from 1 to the number of nucleotides in the sequence, less the length of the probe - 1. Longer probes that hybridize to *the copB* gene trader low, medium, medium-high and high stringency conditions are also contemplated, including those that comprise the entire nucleotide sequence of TTA24 (SEQ ID NO:34), 012E (SEQ ID NO:28) or 046E (SEQ ID NO:32). This hypothetical may be repeated for probes having lengths of about 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80, 90, 100 and greater bases.

In that the CopB proteins and antigenic epitopes of the present invention are believed to be indicative of pathogenic *Moraxella* species, the probes of the present

invention will find particular utility as the basis for diagnostic hybridization assays for detecting *copB* genes in clinical samples. Exemplary clinical samples that can be used in the diagnosis of infections are thus any samples which could possibly include *Moraxella* nucleic acid, including middle ear fluid, sputum, mucus, bronchoalveolar fluid, amniotic fluid or the like. A variety of hybridization techniques and systems are known which can be used in connection with the hybridization aspects of the invention, including diagnostic assays such as those described in Falkow *et al.*, U.S. Patent 4,358,535. Depending on the application envisioned, one will desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of the probe toward the target sequence. For applications requiring a high degree of selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, for example, one will select relatively low salt and/or high temperature conditions, such as provided by 0.02M-0.15M NaCl at temperatures of 50°C to 70°C. These conditions are particularly selective, and tolerate little, if any, mismatch between the probe and the template or target strand.

Of course, for some applications, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template (discussed above), less stringent hybridization conditions are called for in order to allow formation of the heteroduplex. In these circumstances, one would desire to employ conditions such as 0. 15M-0.9M salt, at temperatures ranging from 20°C to 55°C. In any case, it is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide, which serves to destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions can be readily manipulated, and the method of choice will generally depend on the desired results.

In certain embodiments, one may desire to employ nucleic acid probes to isolate variants from clone banks containing mutated clones. In particular embodiments, mutant clone colonies growing on solid media which contain variants of the *copB* nucleic acid sequence could be identified on duplicate filters using hybridization conditions and methods, such as those used in colony blot assays, to obtain hybridization only between probes containing sequence variants and nucleic acid sequence variants contained in

specific colonies. In this manner, small hybridization probes containing short variant sequences of the *copB* gene may be utilized to identify those clones growing on solid media which contain sequence variants of the entire *copB* gene. These clones can then be grown to obtain desired quantities of the variant *copB* nucleic acid sequences or the corresponding CopB antigen.

In clinical diagnostic embodiments, nucleic acid sequences of the present invention are used in combination with an appropriate means, such as a label, for determining hybridization. A wide variety of appropriate indicator means are known in the art, including radioactive, enzymatic or other ligands, such as avidin/biotin, which are capable of giving a detectable signal. In preferred diagnostic embodiments, one will likely desire to employ an enzyme tag such as urease, alkaline phosphatase or peroxidase, instead of radioactive or other environmental undesirable reagents. In the case of enzyme tags, colorimetric indicator substrates are known which can be employed to provide a means visible to the human eye or spectrophotometrically, to identify specific hybridization with pathogen nucleic acid-containing samples. Biotin/avidin-based reagents also may be employed, where one of these binding partners is attached to the probe.

In general, it is envisioned that the hybridization probes described herein will be useful both as reagents in solution hybridizations as well as in embodiments employing a solid phase. In embodiments involving a solid phase, the test DNA (or RNA) from suspected clinical samples, such as exudates, body fluids (e.g., amniotic fluid, middle ear effusion, bronchoalveolar lavage fluid) or even tissues, is adsorbed or otherwise affixed to a selected matrix or surface. This fixed, single-stranded nucleic acid is then subjected to specific hybridization with selected probes under desired conditions. The selected conditions will depend on the particular circumstances based on the particular criteria required (depending, for example, on the G+C contents, type of target nucleic acid, source of nucleic acid, size of hybridization probe, etc.). Following washing of the hybridized surface so as to remove nonspecifically bound probe molecules, specific hybridization is detected, or even quantified, by means of the label.

The nucleic acid sequences which encode for the CopB antigens, or their variants, may be useful in conjunction with polymerase chain reaction (PCR) methodology to detect M. *catarrhalis*. In general, by applying the PCR technology as set out, e.g., in U.S. Patent 4,60,102, one may utilize various portions of the *copB* gene sequence as oligonucleotide primers for the PCR amplification of a defined portion of a *copB* nucleic acid in a sample. The amplified portion of the *copB* sequence may then be detected by hybridization with a hybridization probe containing a complementary sequence. In this manner, extremely small concentrations of *M. catarrhalis* nucleic acid may detected in a sample utilizing CopB sequences.

4. Vectors, Host Cells and Cultures for Producing CopB Antigens

In another embodiment, nucleic acids of the present invention may be used advantageously to produce recombinant CopB protein. *M. catarrhalis* cells may be grown on agar plates using brain heart infusion broth as a medium. Plates are incubated at 37°C in candle extinction jars. Outer membrane fragments may be subsequently prepared from these cells by the EDTA-based extraction procedure of Murphy *et al.* (1989).

To assist those of skill in carrying out more particular aspects of the invention, recombinant clones bearing DNA segments encoding CopB were deposited with the American Type Culture Collection (ATCC) on August 4, 1992, under the provisions of the Budapest Treaty. In particular, plasmid pMEH 120 (ATCC accession number 75285) bearing a nucleic segment encoding the CopB antigen of strain 035E was deposited in the form of purified plasmid DNA. Further the nucleotide sequences of the *copB* genes from *M. catarrhalis* strains TTA24, 012E and 046E have been submitted to GenBank and assigned the accession numbers U69980 (SEQ ID NO:34), U69981 (SEQ ID NO:28), and U69982 (SEQ ID NO:32), respectively.

The gene *copB*, encoding the CopB protein of strain 035E, was originally cloned in a pBR322-based recombinant plasmid, designated pMEH100. Subsequently, this gene was subcloned in pBluescript for sequencing analysis. This new plasmid, designated pMEH120, is what was deposited with the ATCC. Recombinant plasmid pMEH120 is a

pBluescript II SK+ vector containing an insert of *M. catarrhalis* chromosomal NA approximately 4.5 kb in size, and encodes a protein of about 80 kD that is reactive with monoclonal antibody 10F3. As will be appreciated by those of skill in the art in light of the detailed disclosure set forth herein, that the invention is in no way intended to be limited by the foregoing or other specific embodiments that were deposited with the ATCC.

Though CopB protein may be isolated from cultures of *M. catarrhalis*, it may prove more efficient to generate the protein in other host cell systems. In addition, the synthesis of mutant polypeptides, the nucleic acids for which must be generated outside the *M. catarrhalis* genome, also will likely be produced most effectively in a recombinant system.

In order to express a recombinant CopB polypeptide, it is necessary to provide an *copB* gene in an expression cassette. The expression cassette contains a CopB-encoding nucleic acid trader transcriptional control of a promoter. A "promoter" refers to a DNA sequence recognized by the synthetic machinery of the cell, or introduced synthetic machinery, required to initiate the specific transcription of a gene. The phrase "under transcriptional control" means that the promoter is in the correct location and orientation in relation to the nucleic acid to control RNA polymerase initiation and expression of the gene. Those promoters most commonly used in prokaryotic recombinant DNA construction include the B-lactamase (penicillinase) and lactose promoter systems (Chang *et al.*, 1978; Itakura *et al.*, 1977; Goeddel *et al.*, 1979) and a tryptophan (trp) promoter system (Goeddel *et al.*, 1980; EPO Appl. Publ. No. 0036776). While these are the most commonly used, other microbial promoters have been discovered and utilized, and details concerning their nucleotide sequences have been published, enabling a skilled worker to ligate them functionally with plasmid vectors (EPO Appl. Publ. No. 0036776).

The appropriate expression cassette can be inserted into a commercially available expression vector by standard subcloning techniques. For example, *the E. coli* vectors pUC or pBluescript™ may be used according to the present invention to produce

recombinant CopB polypeptide *in vitro*. The manipulation of these vectors is well known in the art. In general, plasmid vectors containing replicon and control sequences which are derived from species compatible with the host cell are used in connection with these hosts. The vector ordinarily carries a replication site, as well as marking sequences which are capable of providing phenotypic selection in transformed cells. For example, *E. coli* is typically transformed using pBR322, a plasmid derived from an *E. coli* species (Bolivar *et al.*, 1977). pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR plasmid, or other microbial plasmid or phage must also contain, or be modified to contain, promoters which can be used by the microbial organism for expression of its own proteins.

In addition, phage vectors containing replicon and control sequences that are compatible with the host microorganism can be used as a transforming vector in connection with these hosts. For example, the phage lambda GEM™- 11 may be utilized in making recombinant phage vector which can be used to transform host cells, such as E. *coli* LE392.

In one embodiment, the protein is expressed as a fusion protein with β-gal, allowing rapid affinity purification of the protein. Examples of such fusion protein expression systems are the glutathione S-transferase system (Pharmacia, Piscataway, NJ), the maltose binding protein system (NEB, Beverley, MA), the FLAG system (IBI, New Haven, CT), and the 6xHis system (Qiagen, Chatsworth, CA). Some of these fusion systems produce recombinant protein bearing only a small number of additional amino acids, which are unlikely to affect the functional capacity of the recombinant protein. For example, both the FLAG system and the 6xHis system add only short sequences, both of which are known to be poorly antigenic and which do not adversely affect folding of the protein to its native conformation. Other fusion systems produce proteins where it is desirable to excise the fusion partner from the desired protein. In another embodiment, the fusion partner is linked to the recombinant protein by a peptide sequence containing a specific recognition sequence for a protease. Examples of suitable

sequences are those recognized by the Tobacco Etch Virus protease (Life Technologies, Gaithersburg, MD) or Factor Xa (New England Biolabs, Beverley, MA).

E coli is a preferred prokaryotic host. For example, *E coli* strain RR1 is particularly useful. Other microbial strains which may be used include *E coli* strains such *as E coli* LE392, *E. coli* B, *and E coli* X 1776 ,(ATCC No. 31537). The aforementioned strains, as well *as E coli* W3110 (F-, lambda-, prototrophic, ATCC No. 273325), bacilli such as *Bacillus subtilis,* or other enterobacteriaceae such as *Salmonella typhimurium* or *Serratia marcescens,* and various *Pseudomonas* species may be used. These examples are, of course, intended to be illustrative rather than limiting. Recombinant bacterial cells, for example *E coli,* are grown in any of a number of suitable media, for example LB, and the expression of the recombinant polypeptide induced by adding IPTG to the media or switching incubation to a higher temperature. After culturing the bacteria for a further period of between 2 and 24 hours, the cells are collected by centrifugation and washed to remove residual media. The bacterial cells are then lysed, for example, by disruption in a cell homogenizer and centrifuged to separate the dense inclusion bodies and cell membranes from the soluble cell components. This centrifugation can be performed trader conditions whereby the dense inclusion bodies are selectively enriched by incorporation of sugars such as sucrose into the buffer and centrifugation at a selective speed.

If the recombinant protein is expressed in the inclusion bodies, as is the case in many instances, these can be washed in any of several solutions to remove some of the contaminating host proteins, then solubilized in solutions containing high concentrations of urea *(e.g.* 8M) or chaotropic agents such as guanidine hydrochloride in the presence of reducing agents such as β-mercaptoethanol or dithiothreitol (DTT).

Under some circumstances, it may be advantageous to incubate the polypeptide for several hours under conditions suitable for the protein to undergo a refolding process into a conformation which more closely resembles that of the native protein. Such conditions generally include low protein concentrations less than 500 μg/ml, low levels of reducing agent, concentrations of urea less than 2 M and often the presence of reagents

such as a mixture of reduced and oxidized glutathione which facilitate the interchange of disulphide bonds within the protein molecule.

The refolding process can be monitored, for example, by SDS-PAGE or with antibodies which are specific for the native molecule. Following refolding, the protein can then be purified further and separated from the refolding mixture by chromatography on any of several supports including ion exchange resins, gel permeation resins or on a variety of affinity columns.

Alternatively, the expression system may be eukaryotic in origin. For example, in one such embodiment, the expression system used is the insect baculovirus system, and the expression cassette comprises the baculovirus polyhedron promoter. The gene encoding the protein can be manipulated by standard techniques in order to facilitate cloning into the baculovirus vector. A preferred baculovirus vector is the pBlueBac vector (Invitrogen, Sorrento, CA). The vector carrying the *copB* gene is transfected into *Spodoptera afrugiperda (e.g.,* Sf9) cells by standard protocols, and the cells are cultured and processed to produce the recombinant protein.

There are a variety of other eukaryotic vectors that provide a suitable vehicle in which recombinant CopB can be produced. In various embodiments of the invention, the expression construct may comprise a virus or engineered construct derived from a viral genome. The ability of certain viruses to enter cells via receptor-mediated endocytosis and to integrate into host cell genome and express viral genes stably and efficiently have made them attractive candidates for the transfer of foreign genes into mammalian cells (Ridgeway, 1988; Nicolas and Rubenstein, 1988; Baichwal and Sugden, 1986; Temin, 1986). The first viruses used as vectors were DNA viruses including the papovaviruses (simian virus 40, bovine papilloma virus, and polyoma) (Ridgeway, 1988; Baichwal and Sugden, 1986) and adenoviruses (Ridgeway, 1988; Baichwal and Sugden, 1986) and adenoassociated viruses. Retroviruses also are attractive gene transfer vehicles (Nicolas and Rubenstein, 1988; Temin, 1986) as are vaccina virus (Ridgeway, 1988) adeno-associated virus (Ridgeway, 1988) and HSV (Glorioso *et al.*, 1995). Such vectors may

be used to (i) transform cell lines *in vitro* for the purpose of expressing proteins of interest or (ii) to transform cells *in vitro* or *in vivo* to provide protective polypeptides.

With respect to eukaryotic vectors, the term promoter will be used here to refer to a group of transcriptional control modules that are clustered around the initiation site for RNA polymerase II. Much of the thinking about how promoters are organized derives from analyses of several viral promoters, including those for the HSV thymidine kinase (tk) and SV40 early transcription units. These studies, augmented by more recent work, have shown that promoters are composed of discrete functional modules, each consisting of approximately 7-20 bp of DNA, and containing one or more recognition sites for transcriptional activator or repressor proteins.

At least one module in each promoter functions to position the start site for RNA synthesis. The best known example of this is the TATA box, but in some promoters lacking a TATA box, such as the promoter for the mammalian terminal deoxynucleotidyl transferase gene and the promoter for the SV40 late genes, a discrete element overlying the start site itself helps to fix the place of initiation.

Additional promoter elements regulate the frequency of transcriptional initiation. Typically, these are located in the region 30-110 bp upstream of the start site, although a number of promoters have recently been shown to contain functional elements downstream of the start site as well. The spacing between promoter elements frequently is flexible, so that promoter function is preserved when elements are inverted or moved relative to one another. In the tk promoter, the spacing between promoter elements can be increased to 50 bp apart before activity begins to decline. Depending on the promoter, it appears that individual elements can function either co-operatively or independently to activate transcription.

The particular promoter that is employed to control the expression of a nucleic acid is not believed to be critical, so long as it is capable of expressing the nucleic acid in the targeted cell. Thus, where a human cell is targeted, it is preferable to position the nucleic acid coding region adjacent to and under the control of a promoter that is capable of being expressed in a human cell. Generally speaking, such a promoter might include

either a human or viral promoter. Preferred promoters include those derived from HSV, including α4 promoter. Another preferred embodiment is the tetracycline controlled promoter.

In various other embodiments, the human cytomegalovirus (CMV) immediate early gene promoter, the SV40 early promoter and the Rous sarcoma virus long terminal repeat can be used to obtain high-level expression of transgenes. The use of other viral or mammalian cellular or bacterial phage promoters which are well-known in the art to achieve expression of a transgene is contemplated as well, provided that the levels of expression are sufficient for a given purpose. Tables II and III list several elements/promoters which may be employed, in the context of the present invention, to regulate the expression of a transgene. This list is not intended to be exhaustive of all the possible elements involved in the promotion of transgene expression but, merely, to be exemplary thereof.

Enhancers were originally detected as genetic elements that increased transcription from a promoter located at a distant position on the same molecule of DNA. This ability to act over a large distance had little precedent in classic studies of prokaryotic transcriptional regulation. Subsequent work showed that regions of DNA with enhancer activity are organized much like promoters. That is, they are composed of many individual elements, each of which binds to one or more transcriptional proteins.

The basic distinction between enhancers and promoters is operational. An enhancer region as a whole must be able to stimulate transcription at a distance; this need not be true of a promoter region or its component elements. On the other hand, a promoter must have one or more elements that direct initiation of RNA synthesis at a particular site and in a particular orientation, whereas enhancers lack these specificities. Promoters and enhancers are often overlapping and contiguous, often seeming to have a very similar modular organization.

Additionally any promoter/enhancer combination (as per the Eukaryotic Promoter Data Base EPDB) could also be used to drive expression of a transgene. Use of a T3, T7 or SP6 cytoplasmic expression system is another possible embodiment. Eukaryotic cells

can support cytoplasmic transcription from certain bacterial promoters if the appropriate bacterial polymerase is provided, either as part of the delivery complex or as an additional genetic expression construct.

Host cells include eukaryotic microbes, such as yeast cultures may also be used. *Saccharomyces cerevisiae,* or common baker's yeast is the most commonly used among eukaryotic microorganisms, although a number of other strains are commonly available. For expression in *Saccharomyces,* the plasmid YRp7, for example, is commonly used (Stinchcomb *et al.*, 1979; Kingsman *et al.*, 1979; Tschemper *et al.*, 1980). This plasmid already contains the trpl gene which provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example ATCC No. 44076 or PEP4-1 (Jones, 1977). The presence of the *trpl* lesion as a characteristic of the yeast host cell genome then provides an effective environment for detecting transformation by growth in the absence of tryptophan.

Suitable promoting sequences in yeast vectors include the promoters for 3-phosphoglycerate kinase (Hitzeman *et al.*, 1980) or other glycolytic enzymes (Hess *et al.*, 1968; Holland *et al.*, 1978), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase. In constructing suitable expression plasmids, the termination sequences associated with these genes are also ligated into the expression vector 3' of the sequence desired to be expressed to provide polyadenylation of the mRNA and termination. Other promoters, which have the additional advantage of transcription controlled by growth conditions are the promoter region for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, and the aforementioned glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Any plasmid vector containing a yeast-compatible promoter, origin of replication and termination sequences is suitable.

In addition to eukaryotic microorganisms, cultures of cells derived from multicellular organisms may also be used as hosts. In principle, any such cell culture is workable, whether from vertebrate or invertebrate culture. However, interest has been greatest in vertebrate cells, and propagation of vertebrate cells in culture has become a routine procedure in recent years. Examples of such useful host cell lines are VERO and HeLa cells, Chinese hamster ovary (CHO) cell lines, and W138, BHK, COS-7, 293 and MDCK cell lines. Expression vectors for such cells ordinarily include (if necessary) an origin of replication, a promoter located in from of the gene to be expressed, along with any necessary ribosome binding sites, RNA splice sites, polyadenylation site, and transcriptional terminator sequences.

5. Preparation of Antibodies to CopB

Antibodies that react immunologically with CopB polypeptides or peptides derived therefrom may be readily prepared through use of well-known techniques, such as those exemplified in U.S. Patent 4,196,265. Typically, this technique involves immunizing a suitable animal with a selected immunogen composition, *e.g.*, purified or partially purified protein, synthetic protein or fragments thereof, as discussed in the section on vaccines. Animals to be immunized are mammals such as cats, dogs and horses, although there is no limitation other than that the subject be capable of mounting an immune response of some kind. The immunizing composition is administered in a manner effective to stimulate antibody producing cells. Rodents such as mice and rats are preferred animals, however, the use of rabbit, sheep or frog cells is possible. The use of rats may provide certain advantages, but mice are preferred, with the BALB/c mouse being most preferred as the most routinely used animal and one that generally gives a higher percentage of stable fusions.

For generation of monoclonal antibodies (MAbs), following immunization, somatic cells with the potential for producing antibodies, specifically B lymphocytes (B cells), are selected for use in the MAb generating protocol. These cells may be obtained from biopsied spleens, tonsils or lymph nodes, or from a peripheral blood sample. Spleen cells and peripheral blood cells are preferred, the former because they are a rich

sdevi

Printed by HPS Server
for
WEST

Printer: cm1_7e12_gbeeptr
Date: 08/12/02
Time: 06:55:02

Document Listing

Document	Selected Pages	Page Range
WO009806851	133	1 - 133
Total (1)	133	-

source of antibody-producing cells that are in the dividing plasmablast stage, and the latter because peripheral blood is easily accessible. Often, a panel of animals will have been immunized and the spleen of the animal with the highest antibody titer removed. Spleen lymphocytes are obtained by homogenizing the spleen with a syringe. Typically, a spleen from an immunized mouse contains approximately 5×10^7 to 2×10^8 lymphocytes.

The antibody-producing B cells from the immunized animal are then fused with cells of an immortal myeloma cell line, generally one of the same species as the animal that was immunized. Myeloma cell lines suited for use in hybridoma-producing fusion procedures preferably are non-antibody-producing, have high fusion efficiency and enzyme deficiencies that render them incapable of growing in certain selective media which support the growth of only the desired fused cells, called "hybridomas."

Any one of a number of myeloma cells may be used and these are known to those of skill in the art. For example, where the immunized animal is a mouse, one may use P3-X63/Ag8, X63-Ag8.653, NS1/1.Ag 4 1, Sp210-Ag14, FO, NSO/U, MPC-11, MPC11-X45-GTG 1.7 and S194/5XX0 Bul; for rats, one may use R210.RCY3, Y3-Ag 1.2.3, IR983F and 4B210; and U-266, GM1500-GRG2, LICR-LON-HMy2 and UC729-6 are all useful in connection with human cell fusions.

One preferred murine myeloma cell line is the NS-1 myeloma cell line (also termed P3-NS-1-Ag4-1), which is readily available from the NIGMS Human Genetic Mutant Cell Repository by requesting cell line repository number GM3573. Another mouse myeloma cell line that may be used is the 8-azaguanine-resistant mouse murine myeloma SP2/0 non-producer cell line.

Methods for generating hybrids of antibody-producing spleen or lymph node cells and myeloma cells usually comprise mixing somatic cells with myeloma cells in a 2:1 proportion, though the proportion may vary from about 20:1 to about 1:1, respectively, in the presence of an agent or agents (chemical or electrical) that promote the fusion of cell membranes. Fusion methods using Sendai virus have been described by Kohler & Milstein (1975; 1976), and those using polyethylene glycol (PEG), such as 37% (v/v)

PEG, by Gefter *et al.* (1977). The use of electrically induced fusion methods is also appropriate.

Fusion procedures usually produce viable hybrids at low frequencies, about 1×10^{-6} to 1×10^{-8}. This does not pose a problem, however, as the viable, fused hybrids are differentiated from the parental, unfused cells (particularly the unfused myeloma cells that would normally continue to divide indefinitely) by culture in a selective medium. The selective medium generally is one that contains an agent that blocks the *de novo* synthesis of nucleotides in the tissue culture media. Exemplary and preferred agents are aminopterin, methotrexate and azaserine. Aminopterin and methotrexate block *de novo* synthesis of both purines and pyrimidines, whereas azaserine blocks only purine synthesis. Where aminopterin or methotrexate is used, the media is supplemented with hypoxanthine and thymidine as a source of nucleotides (HAT medium). Where azaserine is used, the media is supplemented with hypoxanthine.

The preferred selection medium is HAT. Only cells capable of operating nucleotide salvage pathways are able to survive in HAT medium. The myeloma cells are defective in key enzymes of the salvage pathway, *e.g.*, hypoxanthine phosphoribosyl transferase (HPRT), and they cannot survive. The B cells can operate this pathway, but they have a limited life span in culture and generally die within about two weeks. Therefore, the only cells that can survive in the selective media are those hybrids formed from myeloma and B cells.

This culturing provides a population of hybridomas from which specific hybridomas are selected. Typically, selection of hybridomas is performed by single-clone dilution in microtiter plates, followed by testing the individual clonal supernatants (after about two to three weeks) for the desired reactivity. The assay should be sensitive, simple and rapid, such as radioimmunoassays, enzyme immunoassays, cytotoxicity assays, plaque assays, dot immunobinding assays, and the like.

The selected hybridomas are then serially diluted and cloned into individual antibody-producing cell lines, which clones can then be propagated indefinitely to provide MAbs. The cell lines may be exploited for MAb production in two basic ways.

A sample of the hybridoma can be injected, usually in the peritoneal cavity, into a histocompatible animal of the type that was used to provide the somatic and myeloma cells for the original fusion. The injected animal develops tumors secreting the specific monoclonal antibody produced by the fused cell hybrid. The body fluids of the animal, such as serum or ascites fluid, can then be tapped to provide MAbs in high concentration. The individual cell lines could also be cultured *in vitro*, where the MAbs are naturally secreted into the culture medium from which they can be readily obtained in high concentrations. MAbs produced by either means may be further purified, if desired, using filtration, centrifugation and various chromatographic methods such as HPLC or affinity chromatography.

Monoclonal antibodies of the present invention also include anti-idiotypic antibodies produced by methods well-known in the art. Monoclonal antibodies according to the present invention also may be monoclonal heteroconjugates, *i.e.*, hybrids of two or more antibody molecules. In another embodiment, monoclonal antibodies according to the invention are chimeric monoclonal antibodies. In one approach, the chimeric monoclonal antibody is engineered by cloning recombinant DNA containing the promoter, leader, and variable-region sequences from a mouse antibody producing cell and the constant-region exons from a human antibody gene. The antibody encoded by such a recombinant gene is a mouse-human chimera. Its antibody specificity is determined by the variable region derived from mouse sequences. Its isotype, which is determined by the constant region, is derived from human DNA.

In another embodiment, monoclonal antibodies according to the present invention is a "humanized" monoclonal antibody, produced by techniques well-known in the art. That is, mouse complementary determining regions ("CDRs") are transferred from heavy and light V-chains of the mouse Ig into a human V-domain, followed by the replacement of some human residues in the framework regions of their murine counterparts. "Humanized" monoclonal antibodies in accordance with this invention are especially suitable for use in *in vivo* diagnostic and prophylactic methods for treating *Moraxella* infections.

As stated above, the monoclonal antibodies and fragments thereof according to this invention can be multiplied according to *in vitro* and *in vivo* methods well-known in the art. Multiplication *in vitro* is carried out in suitable culture media such as Dulbecco's modified Eagle medium or RPMI 1640 medium, optionally replenished by a mammalian serum such as fetal calf serum or trace elements and growth-sustaining supplements, *e.g.*, feeder cells, such as normal mouse peritoneal exudate cells, spleen cells, bone marrow macrophages or the like. *In vitro* production provides relatively pure antibody preparations and allows scale-up to give large amounts of the desired antibodies. Techniques for large scale hybridoma cultivation under tissue culture conditions are known in the art and include homogenous suspension culture, *e.g.*, in an airlift reactor or in a continuous stirrer reactor or immobilized or entrapped cell culture.

Large amounts of the monoclonal antibody of the present invention also may be obtained by multiplying hybridoma cells *in vivo*. Cell clones are injected into mammals which are histocompatible with the parent cells, *e.g.*, syngeneic mice, to cause growth of antibody-producing tumors. Optionally, the animals are primed with a hydrocarbon, especially oils such as Pristane (tetramethylpentadecane) prior to injection.

In accordance with the present invention, fragments of the monoclonal antibody of the invention can be obtained from monoclonal antibodies produced as described above, by methods which include digestion with enzymes such as pepsin or papain and/or cleavage of disulfide bonds by chemical reduction. Alternatively, monoclonal antibody fragments encompassed by the present invention can be synthesized using an automated peptide synthesizer, or they may be produced manually using techniques well known in the art.

The monoclonal conjugates of the present invention are prepared by methods known in the art, *e.g.*, by reacting a monoclonal antibody prepared as described above with, for instance, an enzyme in the presence of a coupling agent such as glutaraldehyde or periodate. Conjugates with fluorescein markers are prepared in the presence of these coupling agents, or by reaction with an isothiocyanate. Conjugates with metal chelates are similarly produced. Other moieties to which antibodies may be conjugated include

radionuclides such as ^{3}H, ^{125}I, ^{131}I ^{32}P, ^{35}S, ^{14}C, ^{51}Cr, ^{36}Cl, ^{57}Co, ^{58}Co, ^{59}Fe, ^{75}Se, ^{152}Eu, and ^{99}mTc, are other useful labels which can be conjugated to antibodies. Radioactively labeled monoclonal antibodies of the present invention are produced according to well-known methods in the art. For instance, monoclonal antibodies can be iodinated by contact with sodium or potassium iodide and a chemical oxidizing agent such as sodium hypochlorite, or an enzymatic oxidizing agent, such as lactoperoxidase. Monoclonal antibodies according to the invention may be labeled with technetium-^{99}m by ligand exchange process, for example, by reducing pertechnate with stannous solution, chelating the reduced technetium onto a Sephadex column and applying the antibody to this column or by direct labeling techniques, *e.g.*, by incubating pertechnate, a reducing agent such as $SNCl_2$, a buffer solution such as sodium-potassium phthalate solution, and the antibody.

6. Use of Peptides and Monoclonal Antibodies in Immunoassays

It is proposed that the monoclonal antibodies of the present invention will find useful application in standard immunochemical procedures, such as ELISA and western blot methods, as well as other procedures which may utilize antibodies specific to CopB epitopes. While ELISAs are preferred forms of immunoassays, it will be readily appreciated that assays also include RIAs and other non-enzyme linked antibody binding assays or procedures. Additionally, it is proposed that monoclonal antibodies specific to the particular CopB epitope may be utilized in other useful applications. For example, their use in immunoabsorbent protocols may be useful in purifying native or recombinant CopB species or variants thereof.

It also is proposed that the disclosed CopB peptides of the invention will find use as antigens for raising antibodies and in immunoassays for the detection of anti-CopB antigen-reactive antibodies and CopB antigens (competitive assays). Alternatively, immunoassays may be exploited to determine the antigen relationship between certain CopB mutant peptides. Such screening assays may involve (i) the generation of antisera or antibodies against mutant peptides, the reactivity of which can then be checked against other peptides, or (ii) the testing of mutant peptide reactivity with a battery of

immunoreagents developed against heterologous antigens, such as antibody 10F3. In this way, a mutational analysis of various epitopes may be performed.

Diagnostic immunoassays include direct culturing of bodily fluids, either in liquid culture or on a solid support such as nutrient agar. A typical assay involves collecting a sample of bodily fluid from a patient and, optionally, placing the sample in conditions optimum for growth of the pathogen. A determination can then be made as to whether a particular microbe exists in the sample. Further analysis can be carried out to determine the properties of the microbe, *e.g.*, whether or not it is hemolyzing.

Immunoassays encompassed by the present invention include, but are not limited to those described in U.S. Patent No. 4,367,110 (double monoclonal antibody sandwich assay) and U.S. Patent No. 4,452,901 (western blot). Other assays include immunoprecipitation of labeled ligands and immunocytochemistry, both *in vitro* and *in vivo*.

Immunoassays, in their most simple and direct sense, are binding assays. Certain preferred immunoassays are the various types of enzyme linked immunosorbent assays (ELISAs) and radioimmunoassays (RIAs) known in the art. Immunohistochemical detection using tissue sections is also particularly useful. However, it will be readily appreciated that detection is not limited to such techniques, and western blotting, dot blotting, FACS analyses, and the like may also be used.

In one exemplary ELISA, the anti-CopB antibodies of the invention are immobilized onto a selected surface exhibiting protein affinity, such as a well in a polystyrene microtiter plate. Then, a test composition suspected of containing the desired antigen, such as a clinical sample, is added to the wells. After binding and washing to remove non-specifically bound immune complexes, the bound antigen may be detected. Detection is generally achieved by the addition of another antibody, specific for the desired antigen, that is linked to a detectable label. This type of ELISA is a simple "sandwich ELISA." Detection also may be achieved by the addition of a second antibody specific for the desired antigen, followed by the addition of a third antibody that

has binding affinity for the second antibody, with the third antibody being linked to a detectable label.

In another exemplary ELISA, the samples suspected of containing the CopB antigen are immobilized onto the well surface and then contacted with the anti-CopB antibodies. After binding and appropriate washing, the bound immune complexes are detected. Where the initial antigen specific antibodies are linked to a detectable label, the immune complexes may be detected directly. Again, the immune complexes may be detected using a second antibody that has binding affinity for the first antigen specific antibody, with the second antibody being linked to a detectable label.

Competition ELISAs are also possible in which test samples compete for binding with known amounts of labeled antigens or antibodies. The amount of reactive species in the unknown sample is determined by mixing the sample with the known labeled species before or during incubation with coated wells. (Peptide antigen or antibodies may also be linked to a solid support, such as in the form of beads, dipstick, membrane or column matrix, and the sample to be analyzed applied to the immobilized peptide or antibody.) The presence of reactive species in the sample acts to reduce the amount of labeled species available for binding to the well, and thus reduces the ultimate signal.

Irrespective of the format employed, ELISAs have certain features in common, such as coating, incubating or binding, washing to remove non-specifically bound species, and detecting the bound immune complexes. These are described below.

In coating a plate with either antigen or antibody, one will generally incubate the wells of the plate with a solution of the antigen or antibody, either overnight or for a specified period. The wells of the plate will then be washed to remove incompletely adsorbed material. Any remaining available surfaces of the wells are then "coated" with a nonspecific protein that is antigenically neutral with regard to the test antisera. These include bovine serum albumin (BSA), casein and solutions of milk powder. The coating allows for blocking of nonspecific adsorption sites on the immobilizing surface and thus reduces the background caused by nonspecific binding of antisera onto the surface.

After binding of material to the well, coating with a non-reactive material to reduce background, and washing to remove unbound material, the immobilizing surface is contacted with the antisera or clinical or biological extract to be tested in a manner conducive to immune complex (antigen/antibody) formation. Such conditions preferably include diluting the antisera with diluents such as BSA, bovine gamma globulin (BGG) and phosphate buffered saline (PBS)/Tween. These added agents also tend to assist in the reduction of nonspecific background. The layered antisera is then allowed to incubate for from 2 to 4 hours, at temperatures preferably on the order of 25° to 27°C. Following incubation, the antisera-contacted surface is washed so as to remove non-immunocomplexed material. A preferred washing procedure includes washing with a solution such as PBS/Tween, or borate buffer.

Following formation of specific immunocomplexes between the test sample and the bound antigen or antibody, followed by subsequent washing, the occurrence and even amount of immunocomplex formation may be determined by subjecting same to a second antibody having specificity for the first or for a distinct epitope of the bound antigen. Of course, in a test sample suspected of containing antibodies of human origin, the second antibody will preferably be an antibody having specificity in general for human IgG. To provide a detecting means, the detecting antibody will preferably have an associated enzyme that will generate a color development upon incubating with an appropriate chromogenic substrate. Thus, for example, one will desire to contact and incubate the antisera-bound surface with a urease or peroxidase-conjugated anti-human IgG for a period of time and under conditions which favor the development of immunocomplex formation (*e.g.*, incubation for 2 hours at room temperature in a PBS-containing solution such as PBS-Tween).

After incubation with the enzyme-tagged antibody, and subsequent to washing to remove unbound material, the amount of label is quantified by incubation with a chromogenic substrate such as urea and bromocresol purple or 2,2'-azino-di-(3-ethyl-benzthiazoline-6-sulfonic acid [ABTS] and H_2O_2, in the case of peroxidase as the enzyme label. Quantification is then achieved by measuring the degree of color generation, *e.g.*, using a visible spectra spectrophotometer. Alternatively, the label may

be a chemiluminescent one. The use of such labels is described in U.S. Patent Nos. 5,310,687, 5,238,808 and 5,221,605.

In competitive formats, one may use labeled or unlabeled peptide as a competitor for the antigen in a sample. Alternatively, competing antibodies will require that the binding of two different antibodies be distinguishable. This can be accomplished by labeling species and not the other, or by labeling both species with differential labels (*e.g.*, rhodamine and fluorescein).

7. Use of CopB Peptide and CopB-Specific Antibody *In Vivo*

In a further embodiment of the present invention, there are provided methods for active and passive immunoprophylaxis. Active immunoprophylaxis will be discussed first, followed by a discussion of passive immunoprophylaxis. It should be noted that the discussion of formulating vaccine compositions in the context of active immunoprophylaxis is relevant to raising antibodies in experimental animals for passive immunoprophylaxis and for the generation of diagnostic methods.

A. Active Immunoprophylaxis

According to the present invention, CopB polypeptides or CopB-derived peptides, as discussed above, may be used as vaccine formulations to generate protective anti-*M. catarrhalis* antibody responses *in vivo*. By protective, it is meant only that the immune system of a treated individual is capable of generating a response that reduces, to any extent, the clinical impact of the bacterial infection. This may range from a minimal decrease is bacterial burden to outright prevention of infection. Ideally, the treated subject will not exhibit the more serious clinical manifestations of *M. catarrhalis* infection.

Generally, immunoprophylaxis involves the administration, to a subject at risk, a vaccine composition. In the instant case, the vaccine composition will contain a CopB polypeptide or immunogenic derivative thereof in a pharmaceutically acceptable carrier, diluent or excipient. As stated above, those of skill in the art are able, through a variety of mechanisms, to identify appropriate antigenic characteristics of CopB and, in so

doing, develop vaccines that will achieve certain goals. These goals include (i) generation of immune responses against specific strains of *M. catarrhalis*, (ii) generation of immune responses against certain groups of *M. catarrhalis* strains and (iii) generation of immune responses against all *M. catarrhalis* strains.

The stability and immunogenicity of CopB antigens may vary and, therefore, it may be desirable to couple the antigen to a carrier molecule. Exemplary carriers are keyhole limpet hemocyanin (KLH) and human and bovine serum albumin, myoglobin, β-galactosidase, penicillinase and bacterial toxoids. Those of skill in the art are aware of proper methods by which peptides can be linked to carriers without destroying their immunogenic value. Synthetic carriers such as multi-poly-DL-alanyl-poly-L-lysine and poly-L-lysine also are contemplated. Coupling generally is accomplished through amino-or carboxy-terminal residues of the antigen, thereby affording the peptide or polypeptide the greatest chance of assuming a relatively "native" conformation following coupling.

It also may be desirable to include in the composition any of a number of different substances referred to as adjuvants, which are known to stimulate the appropriate portion of the immune system of the vaccinated animal. Suitable adjuvants for the vaccination of subjects (including experimental animals; see above) include, but are not limited to oil emulsions such as Freund's complete or incomplete adjuvant (not suitable for livestock use), Marcol 52:Montanide 888 (Marcol is a Trademark of Esso, Montanide is a Trademark of SEPPIC, Paris), squalane or squalene, Adjuvant 65 (containing peanut oil, mannide monooleate and aluminum monostearate), mineral gels such as aluminum hydroxide, aluminum phosphate, calcium phosphate and alum, surfactants such as hexadecylamine, octadecylamine, lysolecithin, dimethyldioctadecylammonium bromide, N,N-dioctadecyl-N,N'- bis(2-hydroxyethyl)-propanediamine, methoxyhexadecylglycerol and pluronic polyols, polyanions such as pyran, dextran sulfate, polyacrylic acid and carbopol, peptides and amino acids such as muramyl dipeptide, dimethylglycine, tuftsin and trehalose dimycolate. Agents include synthetic polymers of sugars (Carbopol), emulsion in physiologically acceptable oil

vehicles such as mannide mono-oleate (Aracel A) or emulsion with 20 percent solution of a perfluorocarbon (Fluosol-DA) also may be employed.

The preparation of vaccines which contain peptide sequences as active ingredients is generally well understood in the art, as exemplified by U.S. Patents 4,608,251, 4,601,903, 4,599,231, 4,599,230, 4,596,792 and 4.578,770, all of which are incorporated herein by reference. Typically, such vaccines are prepared as injectables. Liquid solutions are preferred; solid forms suitable for solution or suspension in liquid prior to injection also may be prepared. The preparation further may be emulsified. The active immunogenic ingredient is often mixed with excipients which are pharmaceutically acceptable and compatible with the active ingredient. Suitable excipients are, for example, water, saline, dextrose, glycerol, ethanol, or the like and combinations thereof. In addition, if desired, the vaccine may contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents, or adjuvants which enhance the effectiveness of the vaccines.

The vaccine preparations of the present invention also can be incorporated into non-toxic carriers such as liposomes or other microcarrier substances, or by conjugation to polysaccharides, proteins or polymers or in combination with Quil-A to form "iscoms" (immunostimulating complexes). These complexes can serve to reduce the toxicity of the antigen, delay its clearance from the host and improve the immune response by acting as an adjuvant. Other suitable adjuvants for use in this embodiment of the present invention include INF, IL-2, IL-4, IL-8, IL-12 and other immunostimulatory compounds. Further, conjugates comprising the immunogen together with an integral membrane protein of prokaryotic origin, such as TraT (see PCT/AU87/00107) may prove advantageous.

The vaccines are conventionally administered parenterally, by injection, for example, either subcutaneously or intramuscularly. Additional formulations which are suitable for other modes of administration include suppositories and, in some cases, oral formulations. For suppositories, traditional binders and carriers may include, for example, polyalkalene glycols or triglycerides: such suppositories may be formed from

mixtures containing the active ingredient in the range of 0.5% to 10%, preferably 1-2%. Oral formulations include such normally employed excipients as, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate and the like. These compositions take the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations or powders and contain 10-95% of active ingredient, preferably 25-70%.

The peptides may be formulated into the vaccine as neutral or salt forms. Pharmaceutically acceptable salts, include the acid addition salts (formed with the free amino groups of the peptide) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups may also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine, and the like.

The vaccines are administered in a manner compatible with the dosage formulation, and in such amount as will be prophylactically effective and immunogenic. The quantity to be administered depends on the subject to be treated, including, *e.g.*, the capacity of the individual's immune system to synthesize antibodies, and the degree of protection desired. Precise amounts of active ingredient required to be administered depend on the judgment of the practitioner. However, suitable dosage ranges are of the order of several hundred micrograms active ingredient to one milligram or so per vaccination. Suitable regimes for initial administration and booster shots are also variable, but are typified by an initial administration followed by subsequent inoculations or other administrations.

The manner of application may be varied widely. Any of the conventional methods for administration of a vaccine are applicable. These are believed to include oral application on a solid physiologically acceptable base or in a physiologically acceptable dispersion, parenterally, by injection or the like. The dosage of the vaccine will depend on the route of administration and will vary according to the size of the host.

In many instances, it will be desirable to have multiple administrations of the vaccine, usually not exceeding six vaccinations, more usually not exceeding four vaccinations and preferably one or more, usually at least about three vaccinations. The vaccinations will normally be at from two to twelve week intervals, more usually from three to five week intervals. Periodic boosters at intervals of 1-5 years, usually three years, will be desirable to maintain protective levels of the antibodies. The course of the immunization may be followed by assays for antibodies for the supernatant antigens. The assays may be performed by labeling with conventional labels, such as radionuclides, enzymes, fluorescers, and the like. These techniques are well known and may be found in a wide variety of patents, such as U.S. Patent Nos. 3,791,932, 4,174,384 and 3,949,064, as illustrative of these types of assays.

B. Passive Immunoprophylaxis

Passive immunity is defined, for the purposes of this application, as the transfer to one organism of an immune response effector that was generated in another organism. The classic example of establishing passive immunity is to transfer antibodies produced in one organism into a second, immunologically compatible animal. By "immunologically compatible," it is meant that the antibody can perform at least some of its immune functions in the new host animal. More recently, as a better understanding of cellular immune functions has evolved, it has become possible to accomplish passive immunity by transferring other effectors, such as certain kinds of lymphocytes, including cytotoxic and helper T cells, NK cells and other immune effector cells. The present invention contemplates both of these approaches.

Antibodies, antisera and immune effector cells are raised using standard vaccination regimes in appropriate animals, as discussed above. The primary animal is vaccinated with at least CopB polypeptide or peptide product according to the present invention, with or without an adjuvant, to generate an immune response. The immune response may be monitored, for example, by measurement of the levels of antibodies produced, using standard ELISA methods.

Once an adequate immune response has been generated, immune effector cells can be collected on a regular basis, usually from blood draws. The antibody fraction can be purified from the blood by standard means, *e.g.*, by protein A or protein G chromatography. In an alternative preferred embodiment, monoclonal antibody-producing hybridomas are prepared by standard means (Coligan *et al.*, 1991). Monoclonal antibodies are then prepared from the hybridoma cells by standard means. If the primary host's monoclonal antibodies are not compatible with the animal to be treated, it is possible that genetic engineering of the cells can be employed to modify the antibody to be tolerated by the animal to be treated. In the human context, murine antibodies, for example, may be "humanized" in this fashion.

Antibodies, antisera or immune effector cells, prepared as set forth above, are injected into hosts to provide passive immunity against microbial infection. For example, an antibody composition is prepared by mixing, preferably homogeneously mixing, at least one antibody with at least one pharmaceutically or veterinarally acceptable carrier, diluent, or excipient using standard methods of pharmaceutical or veterinary preparation. The amount of antibody required to produce a single dosage form will vary depending upon the microbial species being vaccinated against, the individual to be treated and the particular mode of administration. The specific dose level for any particular individual will depend upon a variety of factors including the age, body weight, general health, sex, and diet of the individual, time of administration, route of administration, rate of excretion, drug combination and the severity of the microbial infectation.

The antibody composition may be administered intravenously, subcutaneously, intranasally, orally, intramuscularly, vaginally, rectally, topically or via any other desired route. Repeated dosings may be necessary and will vary, for example, depending on the clinical setting, the particular microbe, the condition of the patient and the use of other treatments.

C. Animal Model for Testing Efficacy of Therapies

The evaluation of the functional significance of antibodies to surface antigens of *M. catarrhalis* has been hampered by the lack of a suitable animal model. The relative lack of virulence of this organism for animals rendered identification of an appropriate model system difficult (Doere, 1986). Attempts to use rodents, including chinchillas, to study middle ear infections caused by *M. catarrhalis* were unsuccessful, likely because this organism cannot grow or survive in the middle ear of these hosts (Doyle, 1989).

Murine short-term pulmonary clearance models have now been developed which permit an evaluation of the interaction of bacteria with the lower respiratory tract as well as assessment of pathologic changes in the lungs. This model reproducibly delivers an inoculum of bacteria to a localized peripheral segment of the murine lung. Bacteria multiply within the lung, but are eventually cleared as a result of (i) resident defense mechanisms, (ii) the development of an inflammatory response, and/or (iii) the development of a specific immune response. Using this model, it has been demonstrated that serum IgG antibody can enter the alveolar spaces in the absence of an inflammatory response and enhance pulmonary clearance of non-typable *H. influenzae* (McGehee *et al.*, 1989), a pathogen with a host range and disease spectrum nearly identical to those of *M. catarrhalis.*

The above model has been successfully adapted for use as a model of *M. catarrhalis* infection and prophylaxis (Unhanand *et al.*, 1992; Verghese *et al.*, 1990). Thus, there is a ready model with which one can test the efficacy of a peptide-based vaccine or an antibody, polyclonal or monoclonal, in passive immunoprophylaxis. Demonstration of efficacy for an anti-bacterial prophylaxis in this model would be accepted by those in the field of medical microbiology as predictive of efficacy in humans due to the similarity in pulmonary clearance mechanisms between mice and humans.

8. Examples

The following examples are included to demonstrate preferred embodiments of the invention. It should be appreciated by those of skill in the art that the techniques

disclosed in the examples which follow represent techniques discovered by the inventor to function well in the practice of the invention, and thus can be considered to constitute preferred modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments which are disclosed and still obtain a like or similar result without departing from the spirit and scope of the invention.

EXAMPLE I: Nucleotide Sequence Analysis of Four *copB* Genes from *M. catarrhalis*

To begin to address the issue of antigenic heterogeneity of CopB surface epitopes among *M. catarrhalis* strains, the inventors isolated and sequenced the *copB* genes from clinical isolates of *M. catarrhalis* that either do or do not react with MAb 10F3. Isolates TTA24 and O46E do not bind to MAb 10F3; whereas, the O12E and O35E isolates do bind MAb 10F3. In addition, TTA24 is a serum-resistant strain and can be used in the mouse model for pulmonary clearance. (Helminen *et al.*, 1994)

M. catarrhalis TTA24 chromosomal DNA was partially digested with *PstI* (Gibco BRL, Gaithersburg, MD) and fragments ranging from 6 to 9 kb were purified from a 0.7% agarose gel using the GeneCleanII™ kit (BIO 101 Inc., La Jolla, CA). These fragments were ligated overnight at 16°C into *PstI*-digested and alkaline phosphatase-treated plasmid pBR322 (New England Biolabs Inc., Beverly, MA) using T4 DNA ligase (Gibco BRL). The ligation reaction mixture was transformed into competent *E. coli* RR1 cells.

Using a colony blot method, transformants were transferred to nitrocellulose pads, cross-linked using the UV Stratalinker 1800™ (Stratagene Inc., La Jolla, CA) and hybridized with a ^{32}P-labeled 1.3 kb *PvuII/XbaI copB* gene fragment from *M. catarrhalis* O35E. This gene probe was shown to hybridize with a 7.8 kb *PstI* fragment of *M. catarrhalis* TTA24 chromosomal DNA in a Southern blot analysis. Reactive clones were tested for expression of the CopB protein by western blot analysis of whole cell lysates with polyclonal antiserum raised against *M. catarrhalis* TTA24 outer membrane vesicles. A recombinant plasmid containing a 7.8 kb insert that encoded the CopB protein was

designated pTFA100. This plasmid was purified using an alkaline lysis protocol and digested with *HinfI*.

A 2.6 kb fragment known to hybridize with the 1.3 kb *copB* gene probe and expected to contain the complete *copB* gene open reading frame of *M. catarrhalis* TTA24 was purified from the agarose gel as described above. Using *BglII* linkers (New England Biolabs Inc), this fragment was ligated into a *BamHI* digested and alkaline phosphatase-treated pBluescript II SK+ (Stratagene Inc.) and transformed into competent *E. coli* HB101 cells. Recombinant strains were screened as described above. A recombinant plasmid containing a 2.6 kb insert and encoding the CopB protein in the host strain was designated pTTA150.

The O12E and O46E chromosomal DNAs were extracted from whole cell isolates as follows. One to two loopfuls of cells of each respective isolate was resuspended into 1 ml of lysis buffer (150 mM NaCl, 50mM Tris, pH 8.0, 10 mM EDTA, pH 8.0). Sodium dodecyl sulfate (SDS) and RNase were added to yield final concentrations of 1% SDS and 50 μg/ml RNase. Resuspended cells were incubated for 1 to 10 minutes at 55-65°C until cell lysis. DNAs were extracted from cell lysates with an equal volume of a mixture of phenol, chloroform and isoamyl alcohol. The upper phase of the extract was separated and DNAs were precipitated with two volumes of 100% ethanol. The isolated DNAs were resuspended in 300-500 μl of Tris EDTA (TE) buffer. SDS and NaCl were added to each resuspension for final concentrations of 0.1 % and 50 mM, respectively. One microliter of a 1 mg/ml solution of proteinase K was also added to each resuspension. The DNAs were then incubated for 1 hour at 50-55°C. Each was extracted once with an equal volume of phenol:chloroform mixture and then precipitated and resuspended in TE buffer as before.

An aliquot of isolated O12E or O46E chromosomal DNA was amplified using PCR with the 5' primer 5'-CAAGCCTCATAATCGGAG-3' (SEQ ID NO:9) and the 3' primer 5'-CCTCCAGTGAAATCGAATC-3' (SEQ ID NO:10) in order to obtain DNA fragments containing the *copB* gene. The PCR reactions were cycled as follows: 1 minute 30 seconds at 94°C; 1 minute 30 seconds at 51°C; 1 minute 30 seconds at 72°C.

These steps were repeated 29 times. The reactions were held for 10 minutes at 72°C after the final cycle. The PCR reactions were then electrophoresed in a 0.7% agarose gel and the respective 2.5 kb bands were cut out and stored at -20°C. The respective DNAs in the frozen gel slices were extracted by spinning in a CoStar Spin-X column, followed by an isopropanol precipitation of the column flow through liquid. The DNAs were resuspended in distilled water and quantitated on a UV spectrophotometer.

The 2.5 kb PCR products of O12E and O46E were sequenced using the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit With AmpliTaq DNA Polymerase, FS, in a ABI 373A automated sequencer following the manufacturer's protocol. Sequencing primers were designed from the O35E *copB* DNA sequence or were derived from the individual DNA sequences of O12E or O46E. The chromosomal DNA isolated from TTA24 was sequenced using standard methods by Lark Sequencing Technologies Inc., Houston, TX.

The *copB* DNA gene sequences cloned from TTA24, O12E and O46E are SEQ ID NO:34, SEQ ID NO:28 and SEQ ID NO:32, respectively. The deduced CopB amino acid sequences from TTA24, O12E and O46E are SEQ ID NO:33, SEQ ID NO:27 and SEQ ID NO:31, respectively.

EXAMPLE II: Identification of a CopB Epitope for 10F3 MAb

Using the computerized DNA sequence analysis program MacVector™ (IBI, New Haven, CT) the inventors aligned the CopB amino acid sequences from O35E, O12E, TTA24, and O46E with each other. Comparative analysis revealed that the CopB proteins from O35E and TTA24 are 98% identical to each other, the CopB proteins from O12E and O35E are 92% identical and the CopB proteins from O46E and O35E are 89% identical. Variability between the protein sequences was localized to relatively small regions. Based on the sequence variability between TTA24 and O35E and the hydrophilicity of the variants, five regions were identified as potentially comprising strain-specific epitopic core sequences of the CopB protein.

Variable region 1 comprised about 26 amino acids corresponding to about positions 275-300 of O35E. Region 2 comprised about 19 amino acids corresponding to about positions 318-336 of O35E. Region 3 comprised about 14 amino acids corresponding to about positions 343-356 of O35E. Region 4 comprised about 19 amino acids corresponding to about positions 392-410 of O35E. Region 5 comprised about 11 amino acids corresponding to about positions 445455 of O35E.

Alignment of the CopB protein sequences of O35E and TTA24 with those from O12E and O46E revealed that not only does variation occur in these five regions but other relatively small regions also show variability. But the variation in the other areas either does not correlate with reactivity, or the lack thereof, of the isolates to MAb 10F3 or the hydrophilicity of the varied amino acids suggests that these areas do not contain strain-specific epitopes to MAb 10F3; however, these regions may contain strain-specific epitopes to other antibodies for the CopB protein.

EXAMPLE III: Reactivity of a CopB Epitopic Peptide with MAb 10F3

To test whether or not any of the five variable regions identified above contain epitopes specific to MAb 10F3, the inventors synthesized oligopeptides from the CopB protein of O35E that encompassed the five regions of variability identified above using standard methods.

Using a direct ELISA system as described below, the inventors tested the ability of the synthesized oligopeptides to bind to MAb 10F3. Of the five, only the oligopeptide corresponding to variable region 1 (SEQ ID NO:1), specifically bound to MAb 10F3.

Comparing the CopB protein sequences of O35E and TTA24 in region 1 revealed a limited number of amino acid differences between the two. Oligopeptides containing the variations between the two proteins were synthesized and utilizing the direct ELISA system as before, the ability of each oligopeptide to bind to MAb 10F3 was tested. Only replacement of the asparagine residue at about position 295 in the CopB protein of O35E and corresponding to about position 21 of the peptide from variable region 1, by an

aspartate residue was associated with the loss of MAb 10F3 binding activity. Indicating that region 1 contains a strain-specific epitope for MAb 10F3.

Direct ELISAs were conducted as follows. Flat-bottomed microtiter ELISA plates (Coming Glass Ware, Coming, NY) were coated with outer membrane vesicles of *M. catarrhalis* strain O35E (5 mg) or synthetic peptides (0.3-300 mg) in coating buffer (50 mM sodium carbonate, pH 9.6) with or without 2% (v/v) glutaraldehyde and incubated overnight at 4°C. After washing three times with PBS-Tween (0.05% v/v) containing sodium azide (0.025% w/v) and blocking with PBS-Tween containing 1% (w/v) of bovine serum albumin (BSA), MAb 10F3 or MAb 17C7 (control) were added and plates were incubated for one hour at 37°C. After washing, antibody binding was detected using alkaline-phosphatase-conjugated goat anti-mouse IgG (Organon Teknika Corp., West Chester, PA) with *p*-nitrophenylphosphate (Sigma Co., St. Louis, MO) in 10% (v/v) diethanolamine buffer (pH 9.8) as enzyme substrate. Absorbance was measured at 410 nm using a MR700 Microplate reader (Dynatech Laboratories, Chantilly, VA).

EXAMPLE IV: Generation of Antibodies with a CopB Epitope

The oligopeptide from variable region 1 (SEQ ID NO:1) which was shown to lack cross-reactivity with MAb 10F3 in the preceding example was covalently coupled to keyhole limpet hemocyanin (KLH) using glutaraldehyde.

BALB/c mice were either immunized intraperitoneally with 50 μg of oligopeptide-KLH conjugate in Freund's complete adjuvant and boosted with the same dose of oligopeptide-KLH conjugate in Freund's incomplete adjuvant on day 28 of the test or they were immunized with a peptide-KLH conjugate containing an unrelated and non-cross-reacting peptide from *Borrelia burgdorferi*. Blood was drawn from the mice on day 42 post-primary immunization and antisera was isolated from it.

Western blot analysis was performed by standard methods. Antisera raised against the oligopeptide-KLH conjugate contained high titers of IgG antibodies which specifically react with the *M. catarrhalis* CopB protein compared to controls in a western

blot system. As can be seen in FIG. 1A and FIG. 1B, the antisera bound to both the CopB protein of strain O35E (FIG. 1B, lane 1) and, less apparently, to the CopB protein of strain O12E (FIG. 1B, lane 2). The antisera did not bind to the CopB proteins from strains TTA24 and O46E (FIG. 1B, lanes 3 and 4, respectively).

EXAMPLE V: Further Identification of a CopB Epitope for 10F3 MAb

In this study, MAb 10F3 was shown to be bactericidal against MAb 10F3-reactive strains. Further comparison of the deduced amino acid sequences of the CopB proteins from four strains of *M. catarrhalis* revealed a high degree of identity among these proteins which in turn facilitated mapping of the MAb 10F3-reactive epitope. These data allowed construction of a fusion protein which bound MAb 10F3 and induced the synthesis of antibodies directed against the surface of *M. catarrhalis*.

Materials and Methods

Bacterial strains, plasmids and growth conditions. The bacterial strains and plasmids used in this study are listed in Table IV, below. *M. catarrhalis* strain 035E has been described in detail (Helminen *et al.*, 1993a; 1993b). *M. catarrhalis* strains were routinely cultured at 37°C on Brain-Heart Infusion (BHI) agar plates (Difco Laboratories, Detroit, MI) in an atmosphere of 95% air-5% CO_2 or in BHI broth. The *Escherichia coli* cloning strains RR1, HB101, DH5α and recombinant strains were grown on Lubria-Bertani (LB) medium (Sambrook *et al.*, 1989) supplemented, when necessary, with an appropriate antimicrobial compound.

TABLE IV

Bacterial Strains and Plasmids Used in This Study

Strain or plasmid	Description	Source or Reference
Moraxella catarrhalis		
035E	Wild-type isolate from middle ear fluid	(Helminen *et al.*, 1993a)
012E	Wild-type isolate from middle ear fluid	J. Nelson
046E	Wild-type isolate from middle ear fluid	J. Nelson
TTA24	Wild-type isolate from transtracheal aspirate	S. Berk; (Helminen *et al.*, 1994)
Escherichia coli		
HB101	Host for cloning experiments	(Sambrook *et al.*, 1989)
RR1	Host for cloning experiments	(Sambrook *et al.*, 1989)
DH5α	Host for cloning experiments	(Sambrook *et al.*, 1989)
Plasmids		
pBR322	Cloning vector; Amp^r Tet^r	(Sambrook *et al.*, 1989)
pBluescript II SK+	Cloning vector; Amp^r	Stratagene
pTTA100	pBR322 with a 7.8 kb *Pst*I fragment of *M. catarrhalis* strain TTA24 chromosomal DNA containing the *copB* gene	This study
pTTA150	pBluescript II SK+ with a 2.6 kb *Hinf*I fragment from pTTA100 containing the *copB* gene	This study
pGEX-4T-2	Cloning vector	Pharmacia
pEP10F3	pGEX-4T-2 with a 78-bp insert encoding amino acids 275-300 of the CopB protein of *M. catarrhalis* strain 035E	This study

Outer membrane protein preparations and western blot analysis. Outer membrane vesicles were prepared from BHI broth-grown *M. catarrhalis* strains as described (Murphy and Loeb, 1989). Proteins present in these outer membrane vesicles were resolved by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) and detected by western blot analysis as described (Helminen *et al.*, 1993a).

Monoclonal antibodies (MAbs) and polyclonal antisera. MAb 10F3 is a murine IgG_{2a} antibody reactive with a surface-exposed epitope of the CopB outer membrane protein of the *M. catarrhalis* strains 035E and 012E (Helminen *et al.*, 1993a). This MAb does not bind to the CopB protein expressed by *M. catarrhalis* strains TTA24 and 046E (Helminen *et al.*, 1993a). This MAb was used in the form of hybridoma culture supernatant fluid for ELISA assays, epitope mapping, western blot analysis and in the indirect antibody-accessibility assay. For use in bactericidal assays, MAb 10F3 was purified using protein A-Sepharose CL-4B (Pharmacia, Piscataway, NJ) as described (Ey *et al.*, 1978). MAb 17C7 is a murine IgG antibody directed against the UspA surface antigen of *M. catarrhalis* (Helminen *et al.*, 1994). MAb 3F12 is a murine IgG antibody directed against the UspA surface antigen of *M. catarrhalis* (Helminen *et al.*, 1994). MAb 3F12 is a murine IgG antibody specific for the major outer membrane protein (MOMP) of *H. ducreyi* (Gulig *et al.*, 1982). These latter two MAbs were used as negative controls in the ELISA and, indirect antibody-accessibility assays, respectively. Polyclonal rabbit antiserum raised against *M. catarrhalis* strain TTA24 outer membrane vesicles was described previously (Helminen *et al.*, 1993a).

Cloning and nucleotide sequence analysis of the *copB* gene of *M. catarrhalis* strain TTA24. Chromosomal DNA obtained from this strain was partially digested with *Pst*I (Gibco BRL, Gaithersburg, MD) and then subjected to agarose gel electrophoresis. Fragments (6-9 kb) were purified by using the GeneCleanII kit (BIO 101 Inc., La Jolla, CA); these fragments were ligated overnight at 16°C into *Pst*I-digested and alkaline phosphatase-treated pBR322 (New England Biolabs Inc., Beverly, MA) using T4 DNA ligase (Gibco BRL Inc., Gaithersburg, MD). The ligation reaction mixture was used to transform competent *E. coli* RR1 cells. Colony material from tetracycline-resistant transformants was transferred to nitrocellulose pads and treated with UV light (UV Stratalinker, Stratagene, La Jolla, CA) to bind DNA to the nitrocellulose. These membranes were probed with a ^{32}P-labeled 1.3 kb *Pvu*II-*Xba*I fragment from the *copB* gene of *M. catarrhalis* strain 035E (Helminen *et al.*, 1993a). Clones that hybridized this *copB* gene probe were tested for expression of the CopB protein by western blot analysis.

PCR™-based amplification and nucleotide sequence of the *copB* genes of *M. catarrhalis* strains 012E and 046E. Chromosomal DNA from *M. catarrhalis* strains 012E and 046E was extracted by standard methods and used as template DNA in a PCR™ system together with primers derived from the nucleotide sequence of the *copB* gene from *M. catarrhalis* strain 035E (Helminen *et al.*, 1993a). Using a PTC 100 Programmable Thermal Controller (MJ Research, Inc., Cambridge, MA) and the GeneAmp PCR™ kit (Roche Molecular Systems Inc., Branchburg, NJ), a 2.5 kb DNA fragment containing the *copB* gene was amplified from each strain. Using Spin-X centrifuge tube filters (Corning Costar Corp., Cambridge, MA), these 2.5 kb PCR™ products were extracted from 0.7% agarose gel slices and resuspended in water. The nucleotide sequence of these two PCR™ products was determined by the use of an Applied Biosystems Model 373 DNA sequencer (Applied Biosystems, Inc., Foster City, CA) according to the manufacturer's protocol. The sequencing reactions were performed with the Applied Biosystems PRISM Ready Reaction DiDeoxy Terminator Cycle Sequencing Kit with AmpliTaq DNA Polymerase according to the manufacturer's directions (Applied Biosystems, Inc., Foster City, CA). Sequencing primers were designed from the nucleotide sequence of the *copB* gene from *M. catarrhalis* strain 035E (Helminen *et al.*, 1993a) or from the individual *copB* sequences.

Nucleotide sequence accession numbers. The nucleotide sequences of the *copB* genes from *M. catarrhalis* strains TTA24, 012E and 046E have been submitted to GenBank and assigned the accession numbers U69980 (SEQ ID NO:34), U69981 (SEQ ID NO:28), and U69982 (SEQ ID NO:32), respectively.

Oligopeptide synthesis. Oligopeptides were synthesized on a Symphony Peptide Synthesizer (Rainin Instrument Co., Inc., Woburn, MA). Molecular weight and purity of the peptides were determined by high performance liquid chromatography and mass spectrometry.

ELISA. Flat-bottom microtiter ELISA plates (Corning Glass Ware, Corning, NY) were coated with outer membrane vesicles (5 μg) from *M. catarrhalis* strain 035E or with synthetic peptides (10 μg) in 50 mM sodium carbonate buffer (pH 9.6) with or

without 2% (v/v) fresh glutaraldehyde and incubated overnight at 4°C. After washing three times with PBS-Tween (PBS containing 0.05% [v/v] Tween-20 and 0.025% [w/v] sodium azide) and blocking with PBS-Tween containing 1% (w/v) bovine serum albumin (BSA) for 1 hr, 100µl portions of MAb 10F3 or MAb 17C7 (negative control) or antiserum in dilutions of 1:10 to 1:1000 in blocking buffer were added and the plates were incubated at 37°C for 1 hr. After washing, antibody binding to outer membrane vesicles or peptides was detected by using alkaline phosphatase-conjugated goat anti-mouse IgG (Organon Teknika Corp., West Chester, PA) with p-nitrophenylphosphate (Sigma Co., St. Louis, MO) in 10% (v/v) diethanolamine buffer (pH 9.8) as enzyme substrate. Absorbance was measured at 410 nm using a MR 700 Microplate reader (Dynatech Laboratories, Chantilly, VA).

Fine mapping of the MAb 10F3-reactive epitope. Overlapping synthetic decapeptides that were N-terminally bound to a membrane composed of derivatized cellulose were obtained from Research Genetics Inc., Huntsville, AL. After five washes with PBS-Tween, the membrane was blocked for 1 hr at room temperature in PBS-Tween containing 10% (wt/vol) non-fat dry milk, and subsequently incubated with MAb 10F3 overnight at 4°C. Following three washes with PBS-Tween, the membrane was incubated overnight at 4°C with gentle rocking with 10^6 cpm of radioiodinated (specific activity 2×10^7 cpm/µg protein), affinity-purified goat anti-mouse immunoglobulin. The membrane was then washed as before and exposed to X-ray film (Fuji RX safety film; Fuji Industries, Tokyo, Japan).

Construction of a glutathione-S-transferase (GST) fusion protein expressing the MAb 10F3-reactive epitope. The pGEX-4T-2 expression vector system (Pharmacia Biotech Inc., Piscataway, NJ) was used to construct a fusion protein containing the MAb 10F3-reactive epitope at the C-terminus of GST. Using the oligonucleotide primers 5'-CGGGATCCCTAGATATAGAAAAAGAT-3' (SEQ ID NO:11) and 5'-CCGCTCGAGCTTGCCTCGATATTTGTTATC-3' (SEQ ID NO:12) derived from the *copB* gene sequence of *M. catarrhalis* strain 035E and containing a *Bam*HI or a *Xho*I restriction site at their 5'-end, respectively, a 78 bp-fragment encoding amino acid residues 275-300 of the CopB protein (*i.e.*, peptide R1, SEQ ID NO:1, in FIG. 2) was

amplified from *M. catarrhalis* strain 035E chromosomal DNA by PCR™, digested with *Bam*HI and *Xho*I, and ligated into *Bam*HI- and *Xho*I-digested pGEX-4T-2 using T4 DNA ligase. The ligation product was transformed into competent *E. coli* DH5α and recombinant clones were screened for reactivity with MAb 10F3 using a colony blot assay described elsewhere (Gulig *et al.*, 1982). The plasmid construct present in one of the MAb 10F3-reactive clones was designated pEP10F3. The correct in-frame position and orientation of the 78 bp-insert with respect to the GST open reading frame were confirmed by nucleotide sequence analysis of the relevant region of pEP10F3 using the vector-derived sequencing primers 5'-CAATGTGCCTGGATGCGTTC-3' (SEQ ID NO:13) and 5'-CAGACAAGCTGTGACCGTCTCC-3' (SEQ ID NO:14). Large quantities of this fusion protein, designated GST-26, and of GST alone were produced and purified according to the manufacturer's directions for purification of GST fusion proteins.

Immunization protocol. Groups of five two-month-old female BALB/c mice (Charles River Breeding Laboratories, Wilmington, Mass.) were immunized on day 1 by intraperitoneal injection with 50/μg GST-26 suspended in 0.2 ml of Freund's complete adjuvant (Difco Laboratories, Detroit, MI) that had been diluted 1:1 in PBS. Control animals were immunized with GST only in adjuvant. These animals were given an intraperitoneal injection on day 28 with the same amount of protein in Freund's incomplete adjuvant (Difco Laboratories, Detroit, MI) that had been diluted 1:1 in PBS. Blood for serum preparation was obtained by standard methods on day 42. All procedures involving animals received approval from the Institutional Animal Care and Use Committee; all animals were housed in accordance with guidelines from the United States Department of Agriculture and the Association for Assessment and Accreditation of Laboratory Animal Care International.

Indirect antibody-accessibility assay. An overnight BHI broth culture of *M. catarrhalis* strain 035E was diluted in PBS buffer containing 10% (v/v) fetal bovine serum and 0.025% (w/v) sodium azide to a density of 110 Klett units (ca. 10^9 colony forming units [cfu]/ml) as measured with a Klett-Summerson colorimeter (Klett Manufacturing Co., New York, N.Y.). Identical portions (100 μl) of this suspension

were added to 1 ml of MAb 10F3 or MAb 3F12 hybridoma culture supernatant or to 1 ml of the PBS buffer described above containing antiserum at a dilution of 1:500. After incubation at 4°C with gentle agitation for 1 hr, the bacterial cells were washed once and then resuspended in 1 ml of the buffer solution. Radioiodinated goat anti-mouse immunoglobulin was added and the mixture was incubated for 1 hr at 4°C with gentle agitation. The cells were then washed four times with 1 ml of the buffer solution, resuspended in 500 µl of triple detergent (Gulig *et al.*, 1982) and transferred to a 12 × 75 mm glass tube. The radioactivity present in each sample was determined by using a gamma counter.

Bactericidal assay. Complement-sufficient normal adult human serum was prepared by standard methods. Complement inactivation was achieved by heating the serum for 30 min at 56°C. An *M. catarrhalis* BHI broth culture was grown to early logarithmic phase and diluted in Veronal-buffered saline containing 0.15% (w/v) gelatin (Sigma Chemicals Co., St. Louis, MO) to a concentration of 2.5×10^4 cfu/ml (Helminen *et al.*, 1993b). Portions (100 µl) of this cell suspension were added to 100 µl of native or heat-inactivated normal human serum containing either 15 µg of purified MAb 10F3 or heat-inactivated antiserum in final dilutions of 1:20 to 1:2000. This mixture was then incubated at 37°C. At time 0 and at 60 and 120 min after the start of the assay, 10 µl aliquots were removed, suspended in 75 µl of BHI broth and spread onto prewarmed BHI agar plates which were then incubated overnight to determine the number of cfu in each sample.

Results

Cloning of the copB gene from *M. catarrhalis* strain TTA24. As a first step in localizing the epitope that bound MAb 10F3, the amino acid sequence of the CopB protein from a *M. catarrhalis* strain unreactive with this MAb (*i.e.*, strain TTA24) was determined. A 1.3 kb *Pvu*II-*Xba*I fragment from the *M. catarrhalis* strain 035E *copB* gene (Helminen *et al.*, 1993a) was used to screen a genomic library constructed from *M. catarrhalis* strain TTA24 in the plasmid vector pBR322. Whole cell lysates prepared from recombinant *E. coli* clones that hybridized this DNA probe were probed in western

blot analysis with antiserum raised against outer membrane vesicles from *M. catarrhalis* strain TTA24. One of these recombinant clones which expressed an 80 kDa protein reactive with this antiserum was shown to contain a 7.8 kb *Pst*I DNA insert, as was expected from previous Southern blot-based analysis of the conservation of the *copB* gene among *M. catarrhalis* strains (Helminen *et al.*, 1993a). The recombinant plasmid was designated pTTA100 and a 2.6 kb *Hinf*I fragment from this plasmid that was predicted to contain the entire *copB* gene was subcloned into pBluescript to obtain the recombinant plasmid pTTA150. Nucleotide sequence analysis of both strands of this 2.6 kb DNA fragment confirmed that it contained the complete *copB* gene from strain TTA24.

Nucleotide sequence analysis of *copB* genes from additional *M. catarrhalis* strains. To obtain information about the amino acid sequence of other CopB proteins, the *copB* genes from a MAb 10F3-reactive strain (012E) and a MAb 10F3-unreactive strain (046E) were amplified from the chromosome of each strain by PCR™. The oligonucleotide primers (5'-CAAGCCTCATAATCGGAG-3' (SEQ ID NO:9) and 5'-CCTCCAGTGAAATCGAATC-3' (SEQ ID NO:10) were derived from the sequence of the *copB* gene of strain 035E (Helminen *et al.*, 1993a) and were located just outside the open reading frame. Both strands of these two PCR™ products were sequenced in their entirety.

Comparison of the CopB proteins from MAb 10F3-reactive and -unreactive strains. In an attempt to further localize the epitope for MAb 10F3 within the CopB protein, the deduced amino acid sequences of the CopB proteins from these four strains were again aligned and compared as in Example II. These sequences included the leader peptide, the existence of which was confirmed previously by N-terminal amino acid sequence analysis of the mature CopB protein from *M. catarrhalis* strain 035E (Helminen *et al.*, 1993a). The overall identities among these four proteins were striking, with the level of identity between the two most dissimilar proteins (*i.e.*, those of strains 035E and 046E) being about 89%-91%. Interestingly, the reactivity of these CopB proteins with MAb 10F3 was not indicative of the degree of identity among these proteins. For example, the MAb 10F3-reactive protein from strain 035E was about 98%

identical to the MAb 10F3-unreactive CopB protein from strain TTA24 and only about 92%-94% identical to the MAb 10F3-reactive protein from strain 012E.

The fact that the CopB proteins of strains 035E and TTA24 shared the highest degree of identity but differed in their ability to bind MAb 10F3 provided the necessary information to begin localizing the MAb 10F3-reactive epitope. Sequence deviations between these two CopB proteins were confined to five distinct regions (designated 1-5) located between amino acid residues 280 and 450. Because MAb 10F3 was known to bind to both native and denatured CopB (Helminen *et al.*, 1993a) and thus in all likelihood to a linear epitope, only one of these five regions was expected to be involved in the binding of this MAb.

Mapping of the MAb 10F3-reactive epitope. Oligopeptides spanning the five regions of dissimilarity between the CopB proteins of strains 035E and TTA24 were synthesized; these contained the amino acid sequences from the MAb 10F3-reactive CopB protein of strain 035E. These peptides were designated R1 through R5 and had the following amino acid sequences:

R1, LDIEKDKKKRTDEQLQAELDNKYAGK (SEQ ID NO:1);

R2, DANGKLVADLDRNNPTQRE (SEQ ID NO:15);

R3, NLEWTGKNLGFANE (SEQ ID NO:16);

R4, DSPSNMHVVATGANINFDK (SEQ ID NO:17); and

R5, RPGFQNQEKTD (SEQ ID NO:18).

The ability of these peptides to bind MAb 10F3 was investigated by means of a direct ELISA assay. When sodium carbonate coating buffer was used to affix the five peptides to the microtiter wells, none of them bound MAb 10F3. Variations in the pH and salt concentration of this coating buffer, and variation of the coating temperature did not result in detectable binding of MAb 10F3 to any of these peptides. When 100 µl portions of MAb 10F3 hybridoma culture supernatant were preincubated with various concentrations of the five peptides and then used to probe microtiter wells coated with

outer membrane vesicles of strain 035E, there was no concentration-dependent reduction of MAb 10F3-binding by any of the five peptides. However, when glutaraldehyde was added to a final concentration of 2% to the five peptides dissolved in coating buffer, wells coated with the 26-residue peptide R1 (SEQ ID NO:1; FIG. 2, table at lower right) readily bound MAb 10F3. In contrast, none of the other four peptides bound this MAb. The negative control MAb 17C7 did not bind any of these peptides.

Amino acid sequence differences between strains 035E and TTA24 within region 1 corresponded to the residues 6, 12, 13, 14 and 21 of peptide R1 (SEQ ID NO:1; FIG. 2). Because the amino acids in positions 12, 13 and 14 also varied among the two 16 MAb 10F3-reactive strains 035E and 012E (FIG. 2), these residues were considered less likely to be essential for binding this MAb. The peptides R1A (SEQ ID NO:7) and R1B (SEQ ID NO:3; FIG. 2) were synthesized to determine whether the aspartic acid (D) in position 6 or the asparagine (N) in position 21 were required for binding of MAb 10F3. As shown in FIG. 2, changing N to D in position 21 (peptide R1B; SEQ ID NO:3) resulted in loss of MAb 10F3 reactivity, while replacing D with N in position 6 (peptide R1A; SEQ ID NO:7) did not affect binding of this MAb. The presence of N in position 21 of peptide R1 (SEQ ID NO:1) (corresponding to residue 295 in the CopB protein) thus appeared critical for binding of MAb 10F3. As expected, a peptide (R1-TTA24; SEQ ID NO:4) containing the amino acid sequence from the same region of the CopB protein from *M. catarrhalis* strain TTA24 did not bind this MAb (FIG. 2).

Based on the assumption that sequential B-cell epitopes are typically 5 to 7 amino acids in length (Kabat, 1970; Schechter, 1971), an oligopeptide (R1E; SEQ ID NO:6 in FIG. 2) which contained 5 amino acids of the CopB sequence of strain 035E on either side of the N in position 21 of peptide R1 (SEQ ID NO:1) was expected to bind MAb 10F3. However, peptide R1E (SEQ ID NO:6) exhibited no binding of this MAb (FIG. 2). Therefore, the binding site for MAb 10F3 within R1 (SEQ ID NO:1) was mapped using 8 overlapping decapeptides (FIG. 3; Table V) bound (*via* the N-terminus) to a solid phase support composed of derivatized cellulose. Probing of these 8 decapeptides with MAb 10F3 revealed that peptide 7 (SEQ ID NO: 25) and peptide 8 (SEQ ID NO:26) readily bound MAb 10F3, with peptide 7 (SEQ ID NO: 25) exhibiting the strongest

binding. The sequence LDNKYAGKGY (SEQ ID NO:25) (residues 293-302 of the CopB protein of strain 035E) thus allowed maximal binding of MAb 10F3 compared to the other decapeptides.

TABLE V

Decapeptide # as shown in FIG. 3	Amino Acid Sequence	Sequence Listing Designation
PEPTIDE 1	LDIEKDKKKR	SEQ ID NO:19
PEPTIDE 2	EKDKKKRTDE	SEQ ID NO:20
PEPTIDE 3	KKKRTDEQLQ	SEQ ID NO:21
PEPTIDE 4	RTDEQLQAEL	SEQ ID NO:22
PEPTIDE 5	EQLQAELDNK	SEQ ID NO:23
PEPTIDE 6	QAELDNKYAG	SEQ ID NO:24
PEPTIDE 7	LDNKYAGKGY	SEQ ID NO:25
PEPTIDE 8	KYAGKGYKLG	SEQ ID NO:26

Using the method of Kyte and Doolittle (1982), further analysis of region 1 within the context of the CopB protein of strain 035E revealed that this region is part of a highly hydrophilic domain (FIG. 4). Use of the Chou-Fasman method (1978) for predicting secondary structure indicated that this region likely displays a primarily α-helical structure, while the binding site for MAb 10F3 is predicted to form a β-turn (FIG. 4).

Immunogenicity of the GST-26 fusion protein. To determine whether the peptide sequence of R1 (SEQ ID NO:1) could induce the synthesis of antibodies to CopB that would bind to the surface of *M. catarrhalis*, this 26-residue peptide was expressed in a fusion construct (GST-26) at the C-terminus of the 26 kDa GST protein and used to immunize mice. As expected, the molecular weight of purified GST-26 was approximately 30 kDa (FIG. 5A and FIG. 5B, lane 5) and was strongly reactive with

MAb 10F3 in western blot analysis (FIG. 5A, lane 5), while GST alone was unreactive with this MAb (FIG. 5A, lane 6). GST-26 was also strongly reactive with MAb 10F3 when used as antigen in the direct ELISA system.

The presence of CopB-specific antibody in mouse antiserum raised against GST-26 was assessed by ELISA using the synthetic peptides R1 (SEQ ID NO:1), R1A (SEQ ID NO:7), R1E (SEQ ID NO:6), R1B (SEQ ID NO:3) and R1-TTA24 (SEQ ID NO:4) (see FIG. 2) and by western blot analysis using outer membrane vesicles of *M. catarrhalis* strains 035E, 012E, TTA24 and 046E, as well as GST-26 and GST as antigens. In the ELISA assay, the GST-26 antiserum, but not the GST antiserum, reacted strongly with R1 (SEQ ID NO:1), R1A (SEQ ID NO:7), R1E (SEQ ID NO:6), R1B (SEQ ID NO:3) and R1-TTA24 (SEQ ID NO:4) (see FIG. 2). Interestingly, GST-26 antiserum also reacted with R1-TTA24 (SEQ ID NO:4; FIG. 2). Binding occurred regardless of whether glutaraldehyde was included with the peptides for coating of the microtiter plate wells (described above).

In western blot analysis, both the CopB-specific MAb 10F3 (FIG. 5A, lane 1) and antibodies in the GST-26 antiserum (FIG. 5B, lane 1) bound to an 80 kDa antigen of strain 035E. MAb 10F3 also bound to an 82 kDa antigen from strain 012E (FIG. 5A, lane 2); the CopB protein of strain 012E is known to migrate slightly more slowly in SDS-PAGE than the CopB protein of strain 035E. The GST-26 antiserum did not react with any antigens in the 80-82 kDa range from strains 012E, TTA24, and 046E (FIG. 5B, lanes 2-4, respectively). As expected, antibodies in the GST-26 antiserum bound to both GST-26 and GST (FIG. 5B, lane 5 and 6, respectively).

Binding of GST-26 antibodies to the cell surface of *M catarrhalis.* The MAb 10F3-reactive epitope of CopB has been shown to be expressed on the bacterial surface (Helminen *et al.*, 1993a). Because the GST-26 antiserum was shown to contain antibodies specifically binding to the CopB protein from strain 035E and also to region 1 of strain TTA24 (ELISA with R1-TTA24 in FIG. 2), it was of interest to determine whether antibodies derived from this antiserum would bind to the cell surface of *M. catarrhalis.* The use of the indirect antibody-accessibility assay with whole cells of

M. catarrhalis strains 035E, TTA24, 012E and 046E as antigen demonstrated that, as expected, MAb 10F3 readily bound to the surface of strains 035E and 012E, but not to strains TTA24 and 046E (Table VI). Similarly, antibodies in the GST-26 antiserum exhibited specific binding to whole cells of strain 035E, while antibodies present in the control antiserum raised against GST did not (Table VI).

TABLE VI
Binding of Antibodies to Surface Determinants of
Whole Cells of *M. catarrhalis* Strains

	Binding of secondary ^{125}I-labeled goat anti-mouse IgG to murine primary antibody attached to whole cells of *M. catarrhalis* strains[a]			
Primary antibody	035E	012E	TTA24	046E
GST-26 antiserum	62,712	50,099	40,480	14,223
GST antiserum[b]	6,081	4,293	4,795	16,205
MAb 10F3[c]	107,332	63,830	4,231	3,952
MAb 3F12[d]	3,437	3,817	2,651	3,544

[a]Counts per minute of ^{125}I-labeled goat anti-mouse immunoglobulin bound to murine antibodies attached to the bacterial cell surface, as determined by the indirect antibody-accessibility assay. These results are the means of two independent studies.

[b]Antiserum raised in mice against GST was included as a negative control.

[c]MAb 10F3 is included as a positive control.

[d]MAb 3F12, a murine IgG antibody specific for *a H. ducreyi* major outer membrane protein, is included as a negative control.

As expected from the ELISA data, antibodies in the GST-26 antiserum also reacted with whole cells of strains 012E and TTA24 (Table VI), although no binding of the same antiserum to outer membrane vesicles of these strains was detected in western blot analysis (FIG. 5B, lanes 2 and 3, respectively). High levels of binding of the GST control antiserum to whole cells of strain 046E (Table VI) precluded determination of whether the GST-26 antiserum would bind specifically to this organism.

Bactericidal activity of MAb 10F3 and of GST-26 antiserum. MAb 10F3 was previously shown to enhance pulmonary clearance of *M. catarrhalis* from the lungs of mice passively immunized with this MAb (Helminen *et al.*, 1993a). However, the functional basis for this MAb-accelerated elimination of bacteria from the lower

respiratory tract was not determined. Therefore, the ability of MAb 10F3 to kill *M. catarrhalis* in the presence of human complement was evaluated by performing bactericidal assays with the MAb 10F3-reactive strains 035E and 012E and with the MAb 10F3-unreactive strains TTA24 and 046E. Similarly, the antisera raised against GST-26 and GST (negative control) were tested for their bactericidal activity because GST-26 contained antibodies directed at the surface of strains 035E, 012E and TTA24.

All four *M. catarrhalis* strains were resistant to killing by complement-sufficient normal human serum in the absence of specific antibody (FIGS. 6A-6D, closed triangles). In the presence of normal human serum, strains 035E (FIG. 6A, open squares) and 012E (FIG. 6B, open squares) were readily killed by MAb 10F3. The MAb 10F3-unreactive strains TTA24 and 046E were not killed by MAb 10F3 (FIG. 6C and FIG. 6D, open squares). MAb 10F3 did not kill any of these four strains when heat-inactivated normal human serum was used in place of complement-sufficient normal human serum (FIGS. 6A-6D, closed squares).

The GST-26 antiserum did not exert detectable bactericidal activity against strains 035E, TTA24 and 046E (FIG. 6A, FIG. 6C and FIG. 6D, open circles). However, this same GST-26 antiserum did kill strain 012E (FIG. 6B, open circles). The possibility that a prozone effect had prevented killing of strain 035E by the GST-26 antiserum was addressed by incubating the *M. catarrhalis* strain 035E cells with dilutions of the GST-26 antiserum ranging from 1:20 to 1:2000, none of which showed bactericidal activity. Control antiserum raised against GST did not cause complement-mediated killing of any of these *M. catarrhalis* strains (FIGS. 6A-6D, closed circles).

REFERENCES

The following references, to the extent that they provide exemplary procedural or other details supplementary to those set forth herein, are specifically incorporated herein by reference.

U.S. Patent No. 5,310,687

U.S. Patent No. 5,238,808

U.S. Patent No. 5,221,605

U.S. Patent No. 4,608,251

U.S. Patent No. 4,603,102

U.S. Patent No. 4,601,903

U.S. Patent No. 4,599,231

U.S. Patent No. 4,599,230

U.S. Patent No. 4,596,792

U.S. Patent No. 4,578,770

U.S. Patent No. 4,554,101

U.S. Patent No. 4,452,901

U.S. Patent No. 4,367,110

U.S. Patent No. 4,358,535

U.S. Patent No. 4,196,265

U.S. Patent No. 4,174,384

U.S. Patent No. 3,949,064

U.S. Patent No. 3,791,932

Aebi, Stone, Beucher, Cope, Maciver, Thomas, McCracken Jr., Sparling, Hartsen, "Expression of the CopB outer membrane protein *by Moraxella catarrhalis* is regulated by iron and affects iron acquisition from transferrin and lactoferrin," *Infect. Immun.*, 64:2024-2030, 1996.

Anonymous, "Consensus," *Pediatr. Infect. Dis. J.*, 8:594-597, 1989.

Aspin, Hoberman, McCarty, McLinn, Aronoff, Lang, Arrieta, "Comparative study of the safety and efficacy of clarithromycin and amoxicillin-clavulanate in the treatment of acute otitis media in children," *J. Pediatr.* 125:135-141s, 1994.

Baichwal and Sugden, *In: GENE TRANSFER* Kucherlapati, R., ed. New York: Plenum Press, pp. 117-148, 1986.

Bartos and Murphy, "Comparison of the outer membrane proteins of 50 strains of *Branhamella catarrhalis*," *J. Infect. Dis.*, 158:761-765, 1988.

Berk, *Arch. Intern. Med.*, 150:2254-2257, 1990.

Bolivar *et al.*, *Gene*, 2:95, 1977.

Brutlag *et al.*, *CABIOS*, 6:237-245, 1990.

Campagnari, Shanks, Dyer, "Growth of *Moraxella catarrhalis* with human transferrin and lactoferrin: Expression of iron-repressible proteins without siderophore production," *Infect. Immun.*, 62:4909-4914, 1994.

Catlin, *Clin. Microbiol. Rev.*, 3:293-320, 1990.

Chang et a l., *Nature*, 375: 615, 1978.

Chapman Jr., Musher, Jonsson, Clarridge, Wallace Jr., "Development of bacterial antibody during *Branhamella catarrhalis* infection," *J. Infect. Dis.*, 151:878-882, 1985.

Chen, McMichael, van der Meid, Hahn, Mininni, Cowell, Eldridge, "Evaluation of purified UspA from *Moraxella catarrhalis* as a vaccine in a murine model after active immunization," *Infect. Immun.*, 64:1900-1905, 1996.

Chi, Verghese, Moore, Hamati, Berk, "Antibody response to P-protein in patients with *Branhamella catarrhalis* infections," *Am. J. Med.*, 88(5A):255-275, 1990.

Chou and Fasman, "Prediction of the secondary structure of proteins from their amino acid sequence," *Adv. Enzymol.*, 47:45-148, 1978.

Coligan *et al.* (eds), *In: CURRENT PROTOCOLS IN IMMUNOLOGY*, John Wiley, New York, ch. 2.5, 1991.

Davies and Raesen, "The epidemiology of respiratory tract pathogens in Southern Netherlands," *Eur. Respir. J.*, 1:415-420, 1988.

Del Beccaro, Mendelman, Inglis, Richardson, Duncan, Clausen, Stull, "Bacteriology of acute otitis media: A new perspective," *J. Pediatr.*, 120:81-84, 1992.

Doem, *Diag. Microbiol. Infect. Dis.*, 4:191-201, 1986.

Doern, Miller, Winn, "*Branhamella (Neisseria) catarrhalis* systemic disease in humans," *Arch. Intern. Med.*, 141:1690-1692, 1981.

Doyle *Peditr. Infect. Dis. J.* 8(suppl):S45 -7, 1989.

Duim, van Alphen, Eijk, Jansen, Dankeft, "Antigenic drift of non-encapsulated *Haemophilus influenzae* major outer membrane protein P2 in patients with chronic bronchitis is caused by point mutations," *Mol. Microbiol.*, 11:1181-1189, 1994.

Duim, Vogel, Puijk, Jansen, Meloen, Dankeft, van Alphen, "Fine mapping of outer membrane protein P2 antigenic sites which vary during persistent infection by *Haemophilus influenzae*," *Infect. Immun.*, 64:4673-4679, 1996.

Ey, Prowse, Jenkin, "Isolation of pure IgG1, IgG2a, and IgG2b immunoglobulins from mouse serum using protein A-sepharose," *Immunochemistry*, 15:429-436, 1978.

Faden *et al., Pediatr. Infect. Dis. J.*, 9:623-626, 1990.

Faden *et al., Ann. Otol. Rhinol. Laryngol.*, 100:612-615, 1991.

Faden, Hong, Murphy, "Immune response to outer membrane antigens of *Moraxella catarrhalis* in children with otitis media," *Infect. Immun.*, 60:3824-3829, 1992.

Faden, Hong, Pahade, "Immune response to *Moraxella catarrhalis* in children with otitis media: opsonophagocytosis with antigen-coated latex beads," *Ann. Otol. Rhinol. Laryngol.*, 103:522-524, 1994.

Fetrow & Bryant, *Biotechnology*, 11:479-483, 1993.

Gan, Kusmiesz, Shelton, Nelson, "Comparative evaluation of loracarbef and amoxicillin-clavulanate for acute otitis media," *Antimicrob. Agents Chemother.*, 35:967-971, 1991.

Gefter *et al., Somatic Cell Genet.*, 3:231-236M, 1977.

Glorioso *et al.,Ann. Rev. Microbiol.*, 49:675-710, 1995.

Goeddel *et al., Nature*, 281:544, 1979.

Goeddel *et al., Nucleic Acids Res.*, 8:4057, 1980.

Goldblatt, Seymour, Levinsky, Turner, "An enzyme-linked immunosorbent assay for the determination of human IgG subclass antibodies directed against *Branhamella catarrhalis*," *J. Immunol. Methods*, 128:219-225, 1990a.

Goldblatt, Turner, Levinsky, "*Branhamella catarrhalis:* Antigenic determinants and the development of the IgG subclass response in childhood," *J. Infect. Dis.*, 162:1128-1135, 1990b.

Gulig, Frisch, Hensen, "A set of two monoclonal antibodies specific for the cell surface-exposed 39K major outer membrane protein of *Haemophilus influenzae* type b defines all strains of this pathogen," *Infect. Immun.*, 42:516-524, 1983.

Gulig, McCracken Jr., Frisch, Johnston, Hensen, "Antibody response of infants to cell surface-exposed outer membrane proteins of *Haemophilus influenzae* type b after systemic *Haemophilus* disease," *Infect. Immun.*, 37:82-88, 1982.

Hager, Verghese, Alvarez, Berk, "*Branhamella catarrhalis* respiratory infections," *Rev. Infect. Dis.*, 9:1140-1149, 1987.

Helminen, Beach, Maciver, Jarosik, Hansen, Leinonen, "Human immune response against outer membrane proteins of *Moraxella (Branhamella) catarrhalis* determined by immunoblotting and enzyme immunoassay," *Clin. Diagn. Lab. Immunol.*, 2:35-39, 1995.

Helminen, Maciver, Latimer, Cope, McCracken Jr., Hansen, "A major outer membrane protein of *Moraxella catarrhalis* is a target for antibodies that enhance pulmonary clearance of the pathogen in an animal model," *Infect. Immun.*, 61:2003-2010, 1993a.

Helminen, Maciver, Latimer, Klesney-Tait, Cope, Paris, McCracken Jr., Hartsen, "A large, antigenically conserved protein on the surface of *Moraxella catarrhalis* is a target for protective antibodies," *J. Infect. Dis.*, 170:867-872, 1994.

Helminen, Maciver, Latimer, Lumbley, Cope, McCracken Jr., Hansen, "A mutation affecting expression of a major outer membrane protein *of Moraxella catarrhalis* alters serum resistance and survival of this organism *in vivo*," *J. Infect. Dis.*, 168:1194-1201, 1993b.

Hess *et al.*, J. *Adv. Enzyme Reg.*, 7:149, 1968.

Hitzeman *et al.*, *D. Biol. Chem.*, 255:2073, 1980.

Holland *et al.*, *Biochemistry*, 17:4900, 1978.

Hsiao, Sethi, Murphy, "Outer membrane protein CD of *Branhamella catarrhalis* - Sequence conservation in stains recovered from the human respiratory tract," *Microb. Pathog.*, 19:215-225, 1995.

Itakura *et al.*, *Science*, 198:1056, 1977.

Jameson and Wolf, *Cornput. Appl. Biosci.*, 4(1): 181-186, 1988.

Jones, *Genetics*, 85:12, 1977.

Kabat, "Heterogeneity and structure of antibody-combining sites," *Ann. N.Y. Acad. Sci.*, 169:43-54, 1970.

Kingsman *et al.*, *Gene*, 7:141, 1979.

Kohler & Milstein, *Nature*, 256:495-497, 1975.

Kohler & Milstein, *Eur. D. Immunol.*, 6:511-519, 1976.

Kyte and Doolittle, "A simple method for displaying the hydropathic character of a protein," *J. Mol. Biol.*, 157:105-132, 1982.

Leinonen *et al.*, *D. Inject. Dis.*, 144:570-574, 1981.

Maciver, Unhanand, McCracken Jr., Hansen, "Effect of immunization on pulmonary clearance of *Moraxella catarrhalis* in an animal model," *J. Infect. Dis.*, 168:469-472, 1993.

Malkamaki, Honkanen, Leinonen, Makela, "*Branhamella catarrhalis* as a cause of bacteremic pneumonia," *Scand. J. Infect. Dis.*, 15:125-126, 1983.

Marchant, *Am.* D. *Med.*, 88(Suppl. 5A):15S-19S, 1990.

McGehee, *Am. J.Respir. Cell Mol. Biol.*, 1:201-210, 1989.

McLeod, Ahmad, Capewell, Croughan, Calder, Seaton, "Increase in bronchopulmonary infection due to *Branhamella catarrhalis*," *Br. Med. J. [Clin. Res.]*., 292:1103-1105, 1986.

Melendez & Johnson, *Rev. Infect. Dis.*, 13:428-429, 1990.

Murphy, *Pediat. Infect. Dis.* J., 8:S75-S77, 1989.

Murphy and Bartos, "Surface exposed and antigenically conserved determinants of outer membrane proteins of *Branhamella catarrhalis*," *Infect. Immun.*, 57:2938-2941, 1989.

Murphy and Loeb, "Isolation of the outer membrane of *Branhamella catarrhalis*," *Microb. Pathog.*, 6:159-174, 1989.

Murphy *et al.*, *Am. Jrnl. Med.*, 88:5A-41S-5A-45S, 1990.

Murphy, Kirkham, Lesse, "The major heat-modifiable outer membrane protein CD is highly conserved among strains of *Branhamella catarrhalis*," *Mol. Microbiol.*, 10:87-97, 1993.

Nicolas and Rubenstein, *In: Vectors: A survey of molecular cloning vectors and their uses*, Rodriguez and Denhardt (eds.), Stoneham: Butterworth, pp. 494-513, 1988.

Nicotra, Rivera, Liman, Wallace, "*Branhamella catarrhalis* as a lower respiratory tract pathogen in patients with chronic lung disease," *Arch. Intern. Med.*, 146:890-893, 1986.

Pettersson, Haas, van Wassenaar, van der Ley, Tommassen, "Molecular cloning of FrpB, the 70-kilodalton iron-regulated outer membrane protein of *Neisseria meningitidis*," *Infect. Immun.*, 63:4181-4184, 1995.

Pettersson, Kuipers, Pelzer, Verhagen, Tiesjema, Tommassen, Poolman, "Monoclonal antibodies against the 70-kilodalton iron-regulated protein of *Neisseria meningitidis* are bactericidal and strain specific," *Infect. Immun.*, 58:3036-3041, 1990.

Post, Preston, Aul, Larkins-Pettigrew, Rydquist-White, Anderson, Wadowsky, Reagan, Walker, Kingsley, Magit, Ehrlich, "Molecular analysis of bacterial pathogens in otitis media with effusion," *JAMA*, 273:1598-1604, 1995.

Ridgeway, *In:* Rodriguez RL, Denhardt DT, ed. Vectors: A survey of molecular cloning vectors and their uses. Stoneham: Butterworth, pp. 467-492, 1988.

Sambrook, Fritsch, Maniatis, *In: Molecular cloning - a laboratory manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y, 1989.

Sarubbi *et al., Am. J. Med.*, 88(Suppl. 5A):9S-14S, 1990.

Schechter, "Mapping of the combined sites of antibodies specific to polyalanine chains," *Ann. N.Y. Acad. Sci.*, 190:394-419, 1971.

Schonheyder & Ejlertsen, *Eur. d. Clin. Microbiol. Infect. Dis.*, 8:299-300, 1989.

Sethi, Hill, Murphy, "Serum antibodies to outer membrane proteins (OMPs) of *Moraxella (Branhamella) catarrhalis* in patients with bronchiectasis: Identification of OMP B1 as an important antigen," *Infect. Immun.*, 63:1516-1520, 1995.

Stinchcomb *et al., Nature*, 282:39, 1979.

Temin, *In: Gene Transfer*, Kucherlapati (ed.), New York: Plenum Press, pp. 149-188, 1986.

Tschemper *et al., Gene*, 10:157, 1980.

Unhanand *et al.*, J. *Infect. Dis.*, 165:644-650, 1992.

Vaneechoutte, Verschraegen, Claeys, van den Abeele, "Respiratory tract carrier rates of *Moraxella (Branhamella) catarrhalis* in adults and children and interpretation of the isolation of *M. catarrhalis* from sputum," *J. Clin. Microbiol.*, 28:2674-2680, 1990.

Verghese *et al.*, J. *Infect. Dis.*, 162:1189-92, 1990.

Weinberger *et al., Science*, 228:740-742, 1985.

Wolf et *al., Cornput. Appl. Biosci.*, 4(1):187-191, 1988.

Wright & Wallace, *Semin. Respir. Infect.*, 4:40-46, 1989.

Yang, Myers, Mcguinness, Chong, Kwok, Klein, Harkness, "The major outer membrane protein, CD, extracted from *Moraxella (Branhamella) catarrhalis* is a potential vaccine antigen that induces bactericidal antibodies," *FEMS Immunol. Med. Microbiol.*, 17:187-199, 1997.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Board of Regents, The University of Texas System
(B) STREET: 201 West 7th Street
(C) CITY: Austin
(D) STATE: TX
(E) COUNTRY: USA
(F) POSTAL CODE (ZIP): 78701
(G) TELEPHONE: (512) 418-3000
(H) TELEFAX: (512) 474-7577

(ii) TITLE OF INVENTION: DEFINING EPITOPES OF THE OUTER MEMBRANE PROTEIN COPB OF MORAXELLA CATARRHALIS

(iii) NUMBER OF SEQUENCES: 45

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(v) CURRENT APPLICATION DATA:
APPLICATION NUMBER: Unknown
(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 60/023,832
(B) FILING DATE: 12-AUG-1996

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Leu Asp Ile Glu Lys Asp Lys Lys Lys Arg Thr Asp Glu Gln Leu Gln
1 5 10 15

Ala Glu Leu Asp Asn Lys Tyr Ala Gly Lys
20 25

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Leu Asp Ile Lys Lys Asp Asp Lys Thr Leu Thr Glu Thr Glu Leu Gln
1 5 10 15

Ala Glu Leu Asp Asn Lys Tyr Ala Gly Lys
20 25

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Leu Asp Ile Glu Lys Asp Lys Lys Lys Arg Thr Asp Glu Gln Leu Gln
1 5 10 15

Ala Glu Leu Asp Asp Lys Tyr Ala Gly Lys
20 25

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Leu Asp Ile Glu Lys Asn Lys Lys Lys Arg Thr Glu Ala Glu Leu Gln
1 5 10 15

Ala Glu Leu Asp Asp Lys Tyr Ala Gly Lys
20 25

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Ile Asp Ile Glu Lys Lys Gly Lys Ile Arg Thr Glu Ala Glu Leu Leu
1 5 10 15

Ala Glu Leu Asn Lys Asp Tyr Pro Gly Gln
20 25

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Asp Glu Gln Leu Gln Ala Glu Leu Asp Asn Lys Tyr Ala Gly Lys
1 5 10 15

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Leu Asp Ile Glu Lys Asn Lys Lys Lys Arg Thr Asp Glu Gln Leu Gln
1 5 10 15

Ala Glu Leu Asp Asn Lys Tyr Ala Gly Lys
20 25

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

CAAGCCTCAT AATCGGAG 18

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CCTCCAGTGA AATCGAATC 19

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CGGGATCCCT AGATATAGAA AAAGAT 26

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CCGCTCGAGC TTGCCTCGAT ATTTGTTATC 30

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

CAATGTGCCT GGATGCGTTC 20

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

CAGACAAGCT GTGACCGTCT CC 22

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

Asp Ala Asn Gly Lys Leu Val Ala Asp Leu Asp Arg Asn Asn Pro Thr
1 5 10 15

Gln Arg Glu

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

Asn Leu Glu Trp Thr Gly Lys Asn Leu Gly Phe Ala Asn Glu
1 5 10

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Asp Ser Pro Ser Asn Met His Val Val Ala Thr Gly Ala Asn Ile Asn
1 5 10 15

Phe Asp Lys

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

Arg Pro Gly Phe Gln Asn Gln Glu Lys Thr Asp
1 5 10

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Leu Asp Ile Glu Lys Asp Lys Lys Lys Arg
1 5 10

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

Glu Lys Asp Lys Lys Lys Arg Thr Asp Glu
1 5 10

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

Lys Lys Lys Arg Thr Asp Glu Gln Leu Gln
1 5 10

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

Arg Thr Asp Glu Gln Leu Gln Ala Glu Leu
1 5 10

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

Glu Gln Leu Gln Ala Glu Leu Asp Asn Lys
1 5 10

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

Gln Ala Glu Leu Asp Asn Lys Tyr Ala Gly
1 5 10

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

Leu Asp Asn Lys Tyr Ala Gly Lys Gly Tyr
1 5 10

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

Lys Tyr Ala Gly Lys Gly Tyr Lys Leu Gly
1 5 10

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Leu Asp Ile Glu Lys Asp Lys Lys Lys Arg Thr Asp Glu Gln Leu Gln
1 5 10 15

Ala Glu Leu Asp Asn Lys Tyr Ala Gly Lys Gly Tyr Lys Leu Gly
20 25 30

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 758 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

Met Asn Lys Phe Gln Leu Leu Pro Leu Thr Leu Ala Val Ser Ala Ala
1 5 10 15

Phe Thr Thr Thr Ala Phe Ala Ala Val Ser Gln Pro Lys Val Val Leu
20 25 30

Ala Gly Asp Thr Val Val Ser Asp Arg Gln Gly Ala Lys Ile Lys Thr
35 40 45

Asn Val Val Thr Leu Arg Glu Lys Asp Glu Ser Thr Ala Thr Asp Leu
50 55 60

Arg Gly Leu Leu Gln Asp Glu Pro Ala Ile Gly Phe Gly Gly Gly Asn
65 70 75 80

```
Gly Thr Ser Gln Phe Ile Ser Ile Arg Gly Met Gly His Asn Ala Ile
                85                  90                  95

Asp Leu Lys Ile Asp Asn Ala Tyr Gln Asp Gly Gln Leu His Tyr His
            100                 105                 110

Gln Gly Arg Phe Met Leu Asp Pro Gln Met Val Lys Val Val Ser Val
        115                 120                 125

Gln Lys Gly Ala Gly Phe Ala Ser Ala Gly Ile Gly Ala Thr Asn Gly
    130                 135                 140

Ala Ile Val Thr Lys Thr Leu Asp Ala Asp Glu Leu Leu Arg Asn Ser
145                 150                 155                 160

Asp Lys Asp Tyr Gly Phe Lys Val Gly Ala Gly Leu Ser Thr Asn Lys
                165                 170                 175

Gly His Ser Tyr His Gly Ser Ala Phe Gly Lys Ala Gln Thr Gly Phe
            180                 185                 190

Gly Gln Val Asp Ala Leu Val Ser Tyr Asn Gln Val Asn Asp Ser Asp
        195                 200                 205

Tyr Lys Gly Gly Lys Gly Tyr Thr Asn Leu Leu Gly Asn Asp Val Val
    210                 215                 220

Thr Arg Ser Ala Leu Asp Lys Ser Ser Tyr Leu Val Lys Ala Gly Leu
225                 230                 235                 240

Thr Ala Gly Asp His Arg Phe Val Val Ser His Leu Asn Glu Val His
                245                 250                 255

Lys Gly Ile Arg Gly Val Arg Glu Glu Phe Asp Phe Ala Asn Arg Ala
            260                 265                 270

Leu Thr Leu Asp Ile Lys Lys Asp Asp Lys Thr Leu Thr Glu Thr Glu
        275                 280                 285

Leu Gln Ala Glu Leu Asp Asn Lys Tyr Ala Gly Lys Gly Tyr Lys Leu
    290                 295                 300

Gly Ser Lys Thr Pro Asp Gly Lys Lys Tyr Asn Val Val Asp Ala Asn
305                 310                 315                 320

Gly Lys Leu Val Ala Asp Leu Asp Arg Asn Asn Pro Thr Gln Arg Glu
                325                 330                 335

Thr Tyr Gln Lys Leu Thr Asn Leu Glu Trp Thr Gly Lys Asn Leu Gly
            340                 345                 350

Phe Ala Asn Glu Val Thr Ala Asn Val Tyr Lys Leu Glu His Gly Arg
        355                 360                 365
```

```
Asn Ser Ser Ser Asp Gln Gly Asn Thr Tyr Ile Thr Lys Asp Val Pro
    370                 375                 380

Lys Glu Ile Ile Asp Asn Val Asp Thr Pro Ser Asn Met His Val Val
385                 390                 395                 400

Ala Thr Gly Ala Asn Ile Asn Phe Asp Lys Glu Phe Asn His Ser Pro
                405                 410                 415

Leu Lys Gly Phe Gly Val Asp His Thr Leu Leu Lys Tyr Gly Ile Asn
            420                 425                 430

Tyr Arg His Gln Lys Ala Val Pro Pro Arg Ser Leu Lys Pro Gly Val
        435                 440                 445

Val His Gln Glu Lys Thr Asp Ala Gly Ile Tyr Leu Glu Ala Val Asn
    450                 455                 460

Gln Ile Asn Asp Phe Thr Ile Asn Thr Gly Val Arg Val Asp Arg Phe
465                 470                 475                 480

Asp Phe Lys Ala Met Asp Gly Lys Lys Val Gly Lys Thr Asp Ile Asn
                485                 490                 495

Pro Ser Phe Gly Val Ile Tyr Asp Val Asn Pro Asn Leu Ser Val Ser
            500                 505                 510

Gly Asn Leu Ile Tyr Ala Thr Arg Ser Pro Arg Phe Ala Asp Ala Ile
        515                 520                 525

Leu Ser Arg Gly Tyr Arg Gly Gly Val Ile Ser Ile Asp Asp Asn Ala
    530                 535                 540

Lys Ala Glu Lys Ala Arg Asn Thr Glu Ile Gly Phe Asn Tyr Asn Asn
545                 550                 555                 560

Gly Pro Tyr Thr Ala Phe Gly Ser Tyr Phe Trp Gln Arg Val Asp Asn
                565                 570                 575

Ala Arg Ala Thr Ala Asp Ile Thr Arg His Gly Thr Thr Asp Ala Asn
            580                 585                 590

Gly Lys Pro Ile Lys Val Pro Thr Leu Gly Asn Gln Gly His Gln Thr
        595                 600                 605

Asn Gln Gly Tyr Glu Leu Gly Val Gly Tyr Thr Glu Gly Ala Trp Arg
    610                 615                 620

Ala Arg Ala Gly Val Ala His Ser Lys Pro Thr Met His Asn Val Lys
625                 630                 635                 640

Phe Ser Gly Asn Pro Glu Tyr Ala Val Arg Thr Gly Arg Thr Trp Thr
                645                 650                 655

Ala Asp Val Ala Tyr Arg Leu Pro Asn Pro Ser Val Glu Leu Gly Val
            660                 665                 670
```

Arg His Thr Leu Val Glu Gly Val Asp Ala Lys Asp Thr Ser Ile Ile
675 680 685

Ser Gly Glu Val Ser Lys Leu Asn Arg Glu Gly Tyr Asn Val Ser Asp
690 695 700

Ile Tyr Ala Asn Trp Lys Pro Tyr Gly Asn Asp Lys Val Asn Val Asn
705 710 715 720

Phe Ala Val Asn Asn Val Phe Asn Lys Asn Tyr Arg Pro His Thr Gln
725 730 735

Arg Ala Ser Ile Asp Thr Leu Pro Gly Ala Gly Arg Asp Phe Arg Val
740 745 750

Gly Val Asn Phe Thr Tyr
755

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2451 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

AATCGGAGTT ATTATGAATA AGTTTCAATT ATTACCGCTG ACATTGGCGG TGAGTGCCGC 60

TTTTACAACC ACTGCTTTTG CAGCTGTTAG CCAGCCTAAG GTTGTCTTGG CAGGCGATAC 120

AGTGGTCAGT GATCGCCAAG GTGCAAAAAT TAAAACCAAT GTTGTTACCT TACGAGAAAA 180

AGACGAAAGC ACGGCTACAG ATTTGCGTGG TTTGTTACAA GATGAACCTG CCATCGGTTT 240

TGGCGGTGGT AATGGTACTT CTCAATTTAT CAGCATTCGT GGCATGGGTC ATAATGCCAT 300

TGACCTAAAA ATTGACAACG CTTATCAAGA TGGTCAATTA CACTACCACC AAGGTCGCTT 360

TATGCTAGAC CCCCAGATGG TCAAAGTCGT TTCGGTACAA AAAGGGGCAG GCTTTGCCAG 420

TGCAGGCATT GGGGCAACCA ATGGTGCGAT TGTTACCAAA ACCTTAGATG CTGATGAGCT 480

TTTAAGAAAC AGCGACAAAG ATTATGGCTT TAAAGTTGGT GCAGGCTTAT CAACCAACAA 540

AGGTCATTCT TATCATGGCA GTGCCTTTGG TAAAGCACAG ACAGGATTTG GTCAGGTAGA 600

TGCCCTTGTC TCTTATAATC AAGTAAATGA CAGCGACTAT AAAGGCGGTA AAGGGTACAC 660

CAATCTGTTG GGCAATGATG TGGTTACAAG AAGTGCCTTG GATAAATCCA GTTACCTTGT 720

CAAGGCAGGG CTAACTGCTG GCGATCATCG ATTTGTGGTC AGCCATCTAA ATGAAGTTCA 780

TAAAGGCATT CGTGGTGTGC GTGAAGAGTT TGACTTCGCC AATCGTGCCT TGACGCTAGA 840

TATAAAAAAA GATGATAAAA CCCTTACCGA AACTGAGCTT CAGGCAGAGT TAGATAACAA 900

ATATGCAGGC AAGGGTTACA AACTTGGCAG TAAAACACCA GATGGTAAAA AGTATAATGT 960

GGTTGATGCC AATGGTAAAT TGGTGGCTGA TTTAGATAGG AACAACCCAA CTCAGCGTGA 1020

AACCTACCAG AAGTTAACTA ACCTTGAATG GACAGGTAAA AACCTTGGTT TTGCAAATGA 1080

AGTTACTGCC AATGTCTATA AGTTAGAACA TGGACGCAAC TCTTCTAGCG ATCAAGGTAA 1140

CACCTATATC ACCAAAGATG TGCCCAAAGA GATAATAGAT AATGTTGATA CACCATCAAA 1200

CATGCATGTG GTAGCCACAG GGGCTAATAT TAATTTTGAT AAAGAATTTA ATCACAGCCC 1260

ATTAAAAGGC TTTGGTGTTG ACCATACTTT ATTAAAATAT GGCATCAACT ATCGCCATCA 1320

AAAAGCTGTA CCGCCCAGAA GTCTAAAACC TGGTGTGGTG CATCAAGAAA AAACCGATGC 1380

TGGCATTTAT CTAGAAGCGG TTAACCAAAT TAATGACTTT ACCATCAATA CAGGCGTGCG 1440

TGTTGACCGT TTTGACTTTA AAGCCATGGA TGGTAAAAAG GTTGGAAAAA CCGACATCAA 1500

CCCAAGCTTT GGGGTGATTT ATGATGTCAA TCCTAATCTT AGCGTCAGCG GTAACCTAAT 1560

CTATGCCACT CGTAGCCCAC GCTTTGCTGA TGCTATCCTA AGCCGTGGTT ATCGTGGTGG 1620

TGTGATTAGT ATTGATGACA ATGCAAAAGC AGAAAAAGCA CGCAATACCG AGATTGGTTT 1680

TAACTATAAT AATGGACCAT ATACCGCCTT TGGCAGTTAT TTTTGGCAGC GTGTGGATAA 1740

TGCCAGAGCG ACCGCTGATA TAACTCGTCA TGGTACAACA GATGCTAATG GTAAGCCTAT 1800

TAAAGTACCA ACGCTTGGCA ACCAAGGTCA TCAGACCAAC CAAGGCTATG AGCTGGGTGT 1860

AGGCTATACC GAAGGTGCGT GGCGTGCGCG TGCTGGCGTT GCTCACAGCA AACCAACCAT 1920

GCACAATGTC AAATTTAGTG GCAATCCTGA ATATGCCGTG CGTACAGGTC GTACATGGAC 1980

AGCAGATGTC GCCTATCGCC TGCCAAACCC CAGTGTAGAG CTTGGTGTGA GACACACATT 2040

GGTTGAAGGG GTAGATGCCA AAGACACATC CATCATAAGT GGTGAAGTTA GCAAACTTAA 2100

CCGTGAAGGC TATAATGTCA GTGACATCTA TGCCAACTGG AAGCCTTATG GTAATGATAA 2160

GGTGAATGTA AACTTTGCGG TGAATAATGT CTTTAATAAA AACTATCGCC CACACACTCA 2220

GCGTGCTTCC ATAGATACCT TACCTGGGGC AGGTCGTGAT TTCCGTGTTG GCGTGAACTT 2280

CACTTACTAA TACTTACCGA TTTATCGGTA TAATACTGAA CACTCAAGCA CTGCTTGGGT 2340

GTTCTTTTTA TGGGTATGAG TGGATAAAAA CGATAAAAAA CCAATCGTAT CATATTGATT 2400

GGCTATAATG ATAAAATTAA ATCATTACTG ATTTAAAAAC CACGCATTGG C 2451

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 759 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

```
Met Asn Lys Phe Gln Leu Leu Pro Leu Thr Leu Ala Val Ser Ala Ala
1               5                  10                  15

Phe Thr Thr Thr Ala Phe Ala Ala Val Ser Gln Pro Lys Val Val Leu
            20                  25                  30

Ala Gly Asp Thr Val Val Ser Asp Arg Gln Gly Ala Lys Ile Lys Thr
        35                  40                  45

Asn Val Val Thr Leu Arg Glu Lys Asp Glu Ser Thr Ala Thr Asp Leu
    50                  55                  60

Arg Gly Leu Leu Gln Asp Glu Pro Ala Ile Gly Phe Gly Gly Gly Asn
65                  70                  75                  80

Gly Thr Ser Gln Phe Ile Ser Ile Arg Gly Met Gly His Asn Ala Ile
                85                  90                  95

Asp Leu Lys Ile Asp Asn Ala Tyr Gln Asp Gly Gln Leu His Tyr His
            100                 105                 110

Gln Gly Arg Phe Met Leu Asp Pro Gln Met Val Lys Val Val Ser Val
        115                 120                 125

Gln Lys Gly Ala Gly Phe Ala Ser Ala Gly Ile Gly Ala Thr Asn Gly
    130                 135                 140

Ala Ile Val Thr Lys Thr Leu Asp Ala Asp Glu Leu Leu Arg Asn Ser
145                 150                 155                 160

Asp Lys Asp Tyr Gly Phe Lys Val Gly Ala Gly Leu Ser Thr Asn Lys
                165                 170                 175

Gly His Ser Tyr His Gly Ser Ala Phe Gly Lys Ala Gln Thr Gly Phe
            180                 185                 190

Gly Gln Val Asp Ala Leu Val Ser Tyr Asn Gln Val Asn Asp Ser Asp
        195                 200                 205

Tyr Lys Gly Gly Lys Gly Tyr Thr Asn Leu Leu Gly Asn Asp Val Val
    210                 215                 220

Thr Arg Ser Ala Leu Asp Lys Ser Ser Tyr Leu Val Lys Ala Gly Leu
225                 230                 235                 240
```

Thr Ala Gly Asp His Arg Phe Val Val Ser His Leu Asn Glu Val His
245 250 255

Lys Gly Ile Arg Gly Val Arg Glu Glu Phe Asp Phe Ala Asn Arg Ala
260 265 270

Leu Thr Leu Asp Ile Glu Lys Asp Lys Lys Lys Arg Thr Asp Glu Gln
275 280 285

Leu Gln Ala Glu Leu Asp Asn Lys Tyr Ala Gly Lys Gly Tyr Lys Leu
290 295 300

Gly Ser Lys Thr Pro Asp Gly Lys Lys Tyr Asn Val Val Asp Ala Asn
305 310 315 320

Gly Lys Leu Val Ala Asp Leu Asp Arg Asn Asn Pro Thr Gln Arg Glu
325 330 335

Thr Tyr Gln Lys Leu Thr Asn Leu Glu Trp Thr Gly Lys Asn Leu Gly
340 345 350

Phe Ala Asn Glu Val Thr Ala Asn Val Tyr Lys Leu Glu His Gly Arg
355 360 365

Asn Ser Ser Ser Asp Lys Gly Asn Ser Tyr Ile Leu Arg Asp Val Pro
370 375 380

Asn Thr Ile Asn Asp Asn Gly Asp Ser Pro Ser Asn Met His Val Val
385 390 395 400

Ala Thr Gly Ala Asn Ile Asn Phe Asp Lys Glu Phe Asn His Gly Leu
405 410 415

Leu Lys Gly Phe Gly Val Asp His Thr Leu Leu Lys Tyr Gly Ile Asn
420 425 430

Tyr Arg His Gln Glu Ala Val Pro Pro Arg Gly Ile Arg Pro Gly Phe
435 440 445

Gln Asn Gln Glu Lys Thr Asp Ala Gly Ile Tyr Leu Glu Ala Val Asn
450 455 460

Gln Ile Asn Asp Phe Thr Ile Asn Thr Gly Val Arg Val Asp Arg Phe
465 470 475 480

Asp Phe Lys Ala Met Asp Gly Lys Lys Val Gly Lys Thr Asp Ile Asn
485 490 495

Pro Ser Phe Gly Val Ile Tyr Asp Val Asn Pro Asn Leu Ser Val Ser
500 505 510

Gly Asn Leu Ile Tyr Ala Thr Arg Ser Pro Arg Phe Ala Asp Ala Ile
515 520 525

Leu Ser Arg Gly Phe Arg Asp Gly Val Val Ser Ile Ala Asp Asn Ala
530 535 540

Lys Ala Glu Lys Ala Arg Asn Thr Glu Ile Gly Phe Asn Tyr Asn Asn
545 550 555 560

Gly Pro Tyr Thr Ala Phe Gly Ser Tyr Phe Trp Gln Arg Val Asp Asn
565 570 575

Ala Arg Ala Thr Ala Asp Ala Val Gln His Pro Thr Val Thr Thr Ala
580 585 590

Lys Ile Thr Tyr Leu Gly Asn Gln Gly His Gln Thr Asn His Gly Tyr
595 600 605

Glu Leu Gly Val Gly Tyr Thr Glu Gly Ala Trp Arg Ala Arg Ala Gly
610 615 620

Val Ala His Ser Lys Pro Thr Met His Asn Val Lys Phe Lys Ala Asn
625 630 635 640

Pro Glu Tyr Ala Val Arg Thr Gly Arg Thr Trp Thr Ala Asp Val Ala
645 650 655

Tyr Arg Leu Pro Asn Pro Ser Val Glu Leu Gly Val Arg His Thr Leu
660 665 670

Val Glu Gly Val Asp Ala Lys Asp Thr Ser Ile Leu Ser Gly Lys Phe
675 680 685

Asp Asp Lys Asp Gly Ala Ile Leu Asn Arg Glu Gly Tyr Asn Val Ser
690 695 700

Asp Ile Tyr Ala Asn Trp Lys Pro Tyr Gly Asn Asp Lys Val Asn Val
705 710 715 720

Asn Phe Ala Val Asn Asn Val Phe Asn Lys Asn Tyr Arg Pro His Thr
725 730 735

Gln Arg Ala Ser Ile Asp Thr Leu Pro Gly Ala Gly Arg Asp Phe Arg
740 745 750

Val Gly Val Asn Phe Thr Tyr
755

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2407 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

AATCGGAGTT ATTATGAATA AGTTTCAATT ATTACCGCTG ACATTGGCGG TGAGTGCCGC 60

TTTTACAACC ACTGCTTTTG CAGCTGTTAG CCAGCCTAAG GTTGTCTTGG CAGGCGATAC 120

AGTGGTCAGT GATCGCCAAG GTGCAAAAAT TAAAACCAAT GTTGTTACCT TACGAGAAAA 180

AGACGAAAGC ACGGCTACAG ATTTGCGTGG TTTGTTACAA GATGAACCTG CCATCGGTTT 240

TGGCGGTGGT AATGGTACTT CTCAATTTAT CAGCATTCGT GGCATGGGTC ATAATGCCAT 300

TGACCTAAAA ATTGACAACG CTTATCAAGA TGGTCAATTA CACTACCACC AAGGTCGCTT 360

TATGCTAGAC CCCCAGATGG TCAAAGTCGT TTCGGTACAA AAAGGGGCAG GCTTTGCCAG 420

TGCAGGCATT GGGGCAACCA ATGGTGCGAT TGTTACCAAA ACCTTAGATG CTGATGAGCT 480

TTTAAGAAAC AGCGACAAAG ATTATGGCTT TAAAGTTGGT GCAGGCTTAT CAACCAACAA 540

AGGTCATTCT TATCATGGTA GTGCCTTTGG TAAAGCACAG ACAGGATTTG GTCAGGTAGA 600

TGCCCTTGTC TCTTATAATC AAGTAAATGA CAGCGACTAT AAAGGCGGTA AAGGGTACAC 660

CAATCTGTTG GGCAATGATG TGGTTACAAG AAGTGCCTTG GATAAATCCA GTTACCTTGT 720

CAAGGCAGGG CTAACTGCTG GCGATCATCG ATTTGTGGTC AGCCATCTAA ATGAAGTTCA 780

TAAAGGCATT CGTGGCGTGC GTGAAGAGTT TGACTTCGCC AATCGTGCCT TGACGCTAGA 840

TATAGAAAAA GATAAGAAAA AACGTACTGA CGAACAGCTT CAGGCAGAGT TAGATAACAA 900

ATATGCAGGC AAGGGTTACA AACTTGGCAG TAAAACACCA GATGGTAAAA AGTATAATGT 960

GGTTGATGCC AATGGTAAAT TGGTGGCTGA TTTAGATAGG AACAACCCAA CTCAGCGTGA 1020

AACCTACCAA AAGTTAACCA ACCTTGAATG GACAGGTAAA AACCTTGGTT TTGCAAATGA 1080

AGTTACTGCC AATGTCTATA AGTTAGAACA TGGACGCAAC TCCTCTAGCG ATAAAGGTAA 1140

CAGCTATATT CTTCGTGATG TACCTAATAC CATCAATGAT AACGGTGATA GCCCATCAAA 1200

TATGCATGTG GTAGCCACAG GGGCTAATAT TAATTTTGAT AAAGAATTTA ATCACGGTCT 1260

ATTAAAAGGC TTTGGCGTTG ACCATACTTT ATTAAAATAT GGCATCAACT ATCGCCATCA 1320

AGAAGCTGTA CCGCCTAGAG GTATTAGACC TGGTTTTCAA AACCAAGAAA AAACCGATGC 1380

TGGCATTTAT CTAGAAGCGG TTAACCAAAT CAATGACTTT ACCATCAATA CAGGCGTGCG 1440

TGTTGACCGT TTTGACTTTA AAGCTATGGA CGGTAAAAAG GTTGGAAAAA CCGACATCAA 1500

CCCAAGCTTT GGGGTGATTT ATGATGTCAA TCCTAATCTT AGCGTCAGCG GTAACCTAAT 1560

CTATGCCACT CGCAGCCCAC GCTTTGCTGA TGCTATCCTA AGCCGTGGCT TCCGTGATGG 1620

CGTTGTGAGT ATTGCTGATA ACGCAAAAGC AGAAAAAGCA CGCAATACCG AGATTGGTTT 1680

TAACTATAAT AATGGGCCAT ATACCGCCTT TGGCAGCTAT TTTTGGCAGC GTGTGGATAA 1740

TGCCAGAGCT ACTGCCGATG CTGTACAACA CCCCACAGTT ACAACAGCTA AGATTACCTA 1800

TCTTGGCAAC CAAGGTCATC AGACCAACCA CGGTTATGAG CTGGGCGTAG GCTATACCGA 1860

AGGTGCGTGG CGTGCGCGTG CTGGCGTTGC TCACAGCAAG CCAACCATGC ACAATGTCAA 1920

ATTTAAAGCC AACCCTGAAT ATGCCGTGCG TACAGGTCGT ACATGGACAG CAGATGTCGC 1980

CTATCGCCTG CCAAACCCCA GTGTAGAGCT TGGTGTGAGA CACACATTGG TTGAAGGGGT 2040

AGATGCCAAA GACACTTCTA TCCTTAGCGG TAAATTTGAT GATAAAGATG GTGCTATTCT 2100

TAACCGTGAA GGCTATAATG TCAGTGACAT CTATGCCAAC TGGAAGCCTT ATGGCAATGA 2160

TAAGGTGAAT GTAAACTTTG CGGTGAATAA TGTCTTTAAT AAAAACTATC GCCCACACAC 2220

TCAGCGTGCT TCCATAGATA CCTTACCTGG GGCAGGTCGT GATTTCCGTG TTGGCGTGAA 2280

CTTCACTTAC TAATACTTGC CGATTTATCG GTATAATACT GAACACTCAA GCACGCTTGG 2340

GTGTTCTTTT TATGGGTATG AGTGGATAAA AACGATAAAA AAAGCCAATC GTATCATATT 2400

GATTGGC 2407

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 767 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

Met Asn Lys Phe Gln Leu Leu Pro Met Thr Leu Ala Val Ser Ala Ala
1 5 10 15

Phe Thr Thr Thr Ala Phe Ala Ala Val Ser Gln Pro Lys Val Val Leu
20 25 30

Ala Gly Asp Thr Val Val Ser Asp Arg Gln Gly Ala Lys Ile Lys Thr
35 40 45

Asn Val Val Thr Leu Arg Glu Lys Asp Glu Ser Thr Ala Thr Asp Leu
50 55 60

Arg Gly Leu Leu Gln Asp Glu Pro Ala Ile Gly Phe Gly Gly Gly Asn
65 70 75 80

Gly Thr Ser Gln Phe Val Ser Ile Arg Gly Met Gly His Asn Ala Ile
85 90 95

Asp Leu Lys Ile Asp Asn Ala Tyr Gln Asp Gly Gln Leu His Tyr His
100 105 110

```
Gln Gly Arg Phe Met Leu Asp Pro Gln Met Val Lys Val Val Ser Val
        115                 120                 125

Gln Lys Gly Ala Gly Phe Ala Ser Ala Gly Ile Gly Ala Thr Asn Gly
    130                 135                 140

Ala Ile Val Thr Lys Thr Leu Asp Ala Asp Glu Leu Leu Arg Asn Ser
145                 150                 155                 160

Asp Lys Asp Tyr Gly Phe Lys Val Gly Ala Gly Leu Ser Thr Asn Lys
                165                 170                 175

Gly His Ser Tyr His Gly Ser Ala Phe Gly Lys Ala Gln Thr Gly Phe
            180                 185                 190

Gly Gln Val Asp Ala Leu Val Ser Tyr Asn Gln Val Asn Asp Ser Asp
        195                 200                 205

Tyr Lys Gly Gly Lys Gly Tyr Thr Asn Leu Leu Gly Asn Asp Val Val
    210                 215                 220

Thr Arg Ser Ala Leu Asp Lys Ser Ser Tyr Leu Val Lys Ala Gly Leu
225                 230                 235                 240

Thr Ala Gly Asp His Arg Phe Val Val Ser His Leu Asn Glu Val His
                245                 250                 255

Lys Gly Ile Arg Gly Val Arg Glu Glu Phe Asp Phe Ala Asn Arg Ala
            260                 265                 270

Leu Thr Ile Asp Ile Glu Lys Lys Gly Lys Ile Arg Thr Glu Ala Glu
        275                 280                 285

Leu Leu Ala Glu Leu Asn Lys Asp Tyr Pro Gly Gln Gly Tyr Lys Leu
    290                 295                 300

Gly Lys Lys Ile Asn Asp Lys Asp Asn Lys Val Val Gly Tyr His Val
305                 310                 315                 320

Val Asp Ala Asn Gly Arg Gln Val Pro Asn Leu Asn Arg Asn Asn Pro
                325                 330                 335

Thr Gln Arg Glu Thr Tyr Gln Lys Leu Thr Asn Leu Glu Trp Thr Ala
            340                 345                 350

Lys Asn Leu Gly Phe Ala Asn Glu Val Thr Ala Asn Val Tyr Lys Leu
        355                 360                 365

Glu His Gly Arg Asn Ser Ser Ser Asp Gln Gly Asn Thr Tyr Ile Thr
    370                 375                 380

Lys Asp Val Pro Lys Glu Ile Ile Asp Asn Val Asp Thr Pro Ser Asn
385                 390                 395                 400

Met His Val Val Ala Thr Gly Ala Asn Ile Asn Phe Asp Lys Glu Phe
                405                 410                 415
```

```
Asn His Ser Pro Leu Lys Gly Phe Gly Val Asp His Thr Leu Leu Lys
            420                 425                 430

Tyr Gly Ile Asn Tyr Arg His Gln Lys Ala Val Pro Pro Arg Ser Leu
        435                 440                 445

Lys Pro Gly Val Val His Gln Glu Lys Thr Asp Ala Gly Ile Tyr Leu
    450                 455                 460

Glu Ala Val Asn Gln Ile Asn Asp Phe Thr Ile Asn Thr Gly Val Arg
465                 470                 475                 480

Val Asp Arg Phe Asp Phe Lys Ala Leu Gly Gly Lys Lys Val Gly Lys
                485                 490                 495

Thr Asp Ile Asn Pro Ser Phe Gly Val Ile Tyr Asp Val Asn Pro Asn
            500                 505                 510

Leu Ser Val Ser Gly Asn Leu Ile Tyr Ala Thr Arg Ser Pro Arg Phe
        515                 520                 525

Val Asp Ala Ile Leu Ser Arg Gly Tyr Arg Gly Gly Val Ile Ser Ile
    530                 535                 540

Asp Asp Asn Ala Lys Ala Glu Lys Ala Arg Asn Thr Glu Ile Gly Phe
545                 550                 555                 560

Asn Tyr Asn Asn Gly Pro Tyr Thr Ala Phe Gly Ser Tyr Phe Trp Gln
                565                 570                 575

Arg Val Asp Asn Ala Arg Ala Thr Ala Asp Ile Thr Arg His Gly Thr
            580                 585                 590

Thr Asp Ala Asn Gly Lys Ser Ile Lys Val Pro Ala Leu Gly Asn Gln
        595                 600                 605

Gly His Gln Thr Asn Gln Gly Tyr Glu Leu Gly Val Gly Tyr Thr Glu
    610                 615                 620

Gly Ala Trp Arg Ala Arg Ala Gly Val Ala Tyr Ser Lys Pro Thr Met
625                 630                 635                 640

His Asn Val Thr Leu Asp Gly Asn Pro Glu Tyr Ala Val Arg Thr Gly
                645                 650                 655

Arg Thr Trp Thr Ala Asp Val Ala Tyr Arg Leu Pro Asn Pro Ser Val
            660                 665                 670

Glu Leu Gly Val Arg His Thr Leu Val Glu Gly Val Asp Ala Lys Asp
        675                 680                 685

Thr Ser Ile Leu Ser Gly Lys Phe Asp Asp Lys Asp Gly Ala Ile Leu
    690                 695                 700
```

Asn Arg Glu Gly Tyr Asn Val Ser Asp Ile Tyr Ala Asn Trp Lys Pro
705 710 715 720

Tyr Gly Asn Asp Lys Val Asn Val Asn Phe Ala Val Asn Asn Val Phe
725 730 735

Asn Lys Asn Tyr Arg Pro His Thr Gln Arg Ala Ser Ile Asp Thr Leu
740 745 750

Pro Gly Ala Gly Arg Asp Phe Arg Val Gly Val Asn Phe Thr Tyr
755 760 765

(2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2479 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

AATCGGAGTT ATTATGAATA AGTTTCAATT ATTACCGATG ACATTGGCGG TGAGTGCCGC 60

TTTTACAACC ACTGCTTTTG CAGCTGTTAG CCAGCCTAAG GTTGTCTTGG CAGGCGATAC 120

AGTGGTCAGT GATCGCCAAG GTGCAAAAAT TAAAACCAAT GTTGTTACCT TACGAGAAAA 180

AGACGAAAGC ACGGCTACAG ATTTGCGTGG TTTGTTACAA GATGAACCTG CCATCGGTTT 240

TGGCGGTGGT AATGGTACTT CTCAATTTGT CAGCATTCGT GGCATGGGTC ATAATGCCAT 300

TGACCTAAAA ATTGACAACG CTTATCAAGA TGGTCAATTA CACTACCACC AAGGTCGCTT 360

TATGCTAGAC CCCCAGATGG TCAAAGTCGT TTCGGTACAA AAAGGGGCAG GCTTTGCCAG 420

TGCAGGCATT GGGGCAACCA ATGGTGCGAT TGTTACCAAA ACCTTAGATG CTGATGAGCT 480

TTTAAGAAAC AGCGACAAAG ATTATGGCTT TAAAGTTGGT GCAGGCTTAT CAACCAACAA 540

AGGTCATTCT TATCATGGCA GTGCCTTTGG TAAAGCACAG ACAGGATTTG GTCAGGTAGA 600

TGCCCTTGTC TCTTATAATC AAGTAAATGA CAGCGACTAT AAAGGCGGTA AAGGGTACAC 660

CAATCTGTTG GGCAATGATG TGGTTACAAG AAGTGCCTTG GATAAATCCA GTTACCTTGT 720

CAAGGCAGGG CTAACTGCTG GCGATCATCG ATTTGTGGTC AGCCATCTAA ATGAAGTTCA 780

TAAAGGCATT CGTGGTGTGC GTGAAGAGTT TGACTTCGCC AATCGTGCCT TGACAATAGA 840

CATAGAAAAA AAAGGTAAAA TCCGTACCGA AGCAGAGCTT TTGGCAGAAT TAAATAAAGA 900

TTATCCAGGT CAGGGCTATA AGCTTGGTAA AAAAATAAAT GATAAAGATA ATAAGGTGGT 960

CGGTTATCAT GTGGTCGATG CCAATGGCAG ACAGGTACCT AATTTAAATA GGAACAACCC 1020

AACTCAGCGT GAAACCTACC AGAAGTTAAC CAATCTTGAA TGGACAGCTA AAAACCTTGG 1080

TTTTGCAAAT GAAGTTACTG CCAATGTCTA TAAGTTAGAA CATGGACGCA ACTCTTCTAG 1140

CGATCAAGGT AACACCTATA TCACCAAAGA TGTGCCCAAA GAGATAATAG ATAATGTTGA 1200

TACACCATCA AACATGCATG TGGTAGCCAC AGGGGCTAAT ATTAATTTTG ATAAAGAATT 1260

TAATCACAGC CCATTAAAAG GCTTTGGTGT TGACCATACT TTATTAAAAT ATGGCATCAA 1320

CTATCGCCAT CAAAAAGCTG TACCGCCCAG AAGTCTAAAA CCTGGTGTGG TGCATCAAGA 1380

AAAAACCGAT GCTGGCATTT ATCTAGAAGC GGTTAACCAA ATTAATGACT TTACCATCAA 1440

TACAGGCGTG CGTGTTGACC GTTTTGACTT TAAAGCCTTG GGCGGTAAAA AGGTTGGAAA 1500

AACCGACATC AACCCAAGCT TTGGGGTGAT TTATGATGTC AATCCTAATC TTAGCGTCAG 1560

CGGTAACCTA ATCTATGCCA CTCGCAGTCC ACGCTTTGTT GATGCCATCC TAAGCCGTGG 1620

TTATCGTGGT GGTGTGATTA GTATTGATGA CAATGCAAAA GCAGAAAAAG CGCGTAATAC 1680

TGAGATAGGT TTTAACTATA ATAATGGACC ATATACCGCC TTTGGCAGTT ATTTTTGGCA 1740

GCGTGTGGAT AACGCCAGAG CCACCGCTGA TATAACTCGT CATGGTACAA CAGATGCTAA 1800

TGGTAAGTCT ATTAAAGTAC CAGCGCTTGG CAACCAAGGT CATCAGACCA ACCAAGGCTA 1860

TGAGTTGGGC GTAGGTTATA CCGAAGGTGC GTGGCGTGCG CGTGCTGGCG TTGCTTACAG 1920

CAAACCAACC ATGCACAATG TCACATTGGA CGGTAACCCT GAATATGCCG TGCGTACAGG 1980

TCGTACATGG ACAGCAGATG TCGCCTATCG CCTGCCAAAC CCCAGTGTAG AGCTTGGTGT 2040

GAGACACACC TTGGTTGAAG GGGTAGATGC CAAAGACACT TCTATCCTTA GCGGTAAATT 2100

TGATGATAAA GATGGTGCTA TTCTTAACCG TGAAGGCTAT AATGTCAGTG ACATCTATGC 2160

CAACTGGAAG CCTTATGGCA ATGATAAGGT GAATGTAAAC TTTGCGGTGA ATAATGTCTT 2220

TAATAAAAAC TATCGCCCAC ACACTCAGCG TGCTTCCATA GATACCTTAC CTGGGGCAGG 2280

TCGTGATTTC CGTGTTGGCG TGAACTTCAC TTACTAATAC TTACCGATTT ATCGGTATAA 2340

TACTGAACAC TCAAGCACGC TTGGGTGTTC TTTTTATGGG TATGAGTGGA TAAAAACGAT 2400

AAAAAAAGCC AATCGTATCA TATTGATTGG CTATAATGAT AAAATTAAAT CATTACTGAT 2460

TTAAAAACCA CGCATTGGC 2479

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 759 amino acids

(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

```
Met Asn Lys Phe Gln Leu Leu Pro Leu Thr Leu Ala Val Ser Ala Ala
1               5                  10                  15

Phe Thr Thr Thr Ala Phe Ala Ala Val Ser Gln Pro Lys Val Val Leu
            20                  25                  30

Ala Gly Asp Thr Val Val Ser Asp Arg Gln Gly Ala Lys Ile Lys Thr
        35                  40                  45

Asn Val Val Thr Leu Arg Glu Lys Asp Glu Ser Thr Ala Thr Asp Leu
    50                  55                  60

Arg Gly Leu Leu Gln Asp Glu Pro Ala Ile Gly Phe Gly Gly Gly Asn
65                  70                  75                  80

Gly Thr Ser Gln Phe Ile Ser Ile Arg Gly Met Gly His Asn Ala Ile
                85                  90                  95

Asp Leu Lys Ile Asp Asn Ala Tyr Gln Asp Gly Gln Leu His Tyr His
            100                 105                 110

Gln Gly Arg Phe Met Leu Asp Pro Gln Met Val Lys Val Val Ser Val
        115                 120                 125

Gln Lys Gly Ala Gly Phe Ala Ser Ala Gly Ile Gly Ala Thr Asn Gly
    130                 135                 140

Ala Ile Val Thr Lys Thr Leu Asp Ala Asp Glu Leu Leu Arg Asn Ser
145                 150                 155                 160

Asp Lys Asp Tyr Gly Phe Lys Val Gly Ala Gly Leu Ser Thr Asn Lys
                165                 170                 175

Gly His Ser Tyr His Gly Ser Ala Phe Gly Lys Ala Gln Thr Gly Phe
            180                 185                 190

Gly Gln Val Asp Ala Leu Val Ser Tyr Asn Gln Val Asn Asp Ser Asp
        195                 200                 205

Tyr Lys Gly Gly Lys Gly Tyr Thr Asn Leu Leu Gly Asn Asp Val Val
    210                 215                 220

Thr Arg Ser Ala Leu Asp Lys Ser Ser Tyr Leu Val Lys Ala Gly Leu
225                 230                 235                 240

Thr Ala Gly Asp His Arg Phe Val Val Ser His Leu Asn Glu Val His
                245                 250                 255

Lys Gly Ile Arg Gly Val Arg Glu Glu Phe Asp Phe Ala Asn Arg Ala
            260                 265                 270
```

Leu Thr Leu Asp Ile Glu Lys Asn Lys Lys Lys Arg Thr Glu Ala Glu
275 280 285

Leu Gln Ala Glu Leu Asp Asp Lys Tyr Ala Gly Lys Gly Tyr Lys Leu
290 295 300

Gly Ser Lys Thr Pro Asp Gly Lys Lys Tyr Asn Val Val Asp Ala Asn
305 310 315 320

Gly Lys Leu Val Asp Asn Leu Val Gly Asn Asn Pro Thr Gln Arg Glu
325 330 335

Thr Tyr Gln Lys Leu Thr Asn Leu Glu Trp Thr Ala Lys Asn Leu Gly
340 345 350

Phe Ala Asn Glu Val Thr Ala Asn Val Tyr Lys Leu Glu His Gly Arg
355 360 365

Asn Ser Ser Ser Asp Lys Gly Asn Ser Tyr Ile Leu Arg Asp Val Pro
370 375 380

Asn Thr Ile Asn Asp Asn Gly Asp Ser Pro Ser Asn Met His Val Ser
385 390 395 400

Ala Lys Gly Ala Asn Ile Asn Phe Asp Lys Glu Phe Asn His Gly Leu
405 410 415

Leu Lys Gly Phe Gly Val Asp His Thr Leu Leu Lys Tyr Gly Ile Asn
420 425 430

Tyr Arg His Gln Glu Ala Val Pro Pro Arg Gly Ile Arg Pro Gly Phe
435 440 445

Gln His Gln Glu Lys Thr Asp Ala Gly Ile Tyr Leu Glu Ala Val Asn
450 455 460

Gln Ile Asn Asp Phe Thr Ile Asn Thr Gly Val Arg Val Asp Arg Phe
465 470 475 480

Asp Phe Lys Ala Met Asp Gly Lys Lys Val Gly Lys Thr Asp Ile Asn
485 490 495

Pro Ser Phe Gly Val Ile Tyr Asp Val Asn Pro Asn Leu Ser Val Ser
500 505 510

Gly Asn Leu Ile Tyr Ala Thr Arg Ser Pro Arg Phe Ala Asp Ala Ile
515 520 525

Leu Ser Arg Gly Phe Arg Asp Gly Val Val Ser Ile Ala Asp Asn Ala
530 535 540

Lys Ala Glu Lys Ala Arg Asn Thr Glu Ile Gly Phe Asn Tyr Asn Asn
545 550 555 560

Gly Pro Tyr Thr Ala Phe Gly Ser Tyr Phe Trp Gln Arg Val Asp Asn
565 570 575

Ala Arg Ala Thr Ala Asp Ala Val Gln His Pro Thr Val Thr Thr Ala
580 585 590

Lys Ile Thr Tyr Leu Gly Asn Gln Gly His Gln Thr Asn His Gly Tyr
595 600 605

Glu Leu Gly Val Gly Tyr Thr Glu Gly Ala Trp Arg Ala Arg Ala Gly
610 615 620

Val Ala His Ser Lys Pro Thr Met His Asn Val Lys Phe Lys Ala Asn
625 630 635 640

Pro Glu Tyr Ala Val Arg Thr Gly Arg Thr Trp Thr Ala Asp Val Ala
645 650 655

Tyr Arg Leu Pro Asn Pro Ser Val Glu Leu Gly Val Arg His Thr Leu
660 665 670

Val Glu Gly Val Asp Ala Lys Asp Thr Ser Ile Leu Ser Gly Lys Phe
675 680 685

Asp Asp Lys Asp Gly Ala Ile Leu Asn Arg Glu Gly Tyr Asn Val Ser
690 695 700

Asp Ile Tyr Ala Asn Trp Lys Pro Tyr Gly Asn Asp Lys Val Asn Val
705 710 715 720

Asn Phe Ala Val Asn Asn Val Phe Asn Lys Asn Tyr Arg Pro His Thr
725 730 735

Gln Arg Ala Ser Ile Asp Thr Leu Pro Gly Ala Gly Arg Asp Phe Arg
740 745 750

Val Gly Val Asn Phe Thr Tyr
755

(2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2455 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

AATCGGAGTT ATTATGAATA AGTTTCAATT ATTACCGCTG ACATTGGCGG TGAGTGCCGC 60

TTTTACAACC ACTGCTTTTG CAGCTGTTAG CCAGCCTAAG GTTGTCTTGG CAGGCGATAC 120

AGTGGTCAGT GATCGCCAAG GTGCAAAAAT TAAAACCAAT GTTGTTACCT TACGAGAAAA 180

AGACGAAAGC ACGGCTACAG ATTTGCGTGG TTTGTTACAA GATGAACCTG CCATCGGTTT 240

TGGCGGTGGT AATGGTACTT CTCAATTTAT CAGCATTCGT GGCATGGGTC ATAATGCCAT 300

TGACCTAAAA ATTGACAACG CTTATCAAGA TGGTCAATTA CACTACCACC AAGGTCGCTT 360

TATGCTAGAC CCCCAGATGG TCAAAGTCGT TTCGGTACAA AAAGGGGCAG GCTTTGCCAG 420

TGCAGGCATT GGGGCAACCA ATGGTGCGAT TGTTACCAAA ACCTTAGATG CTGATGAGCT 480

TTTAAGAAAC AGCGACAAAG ATTATGGCTT TAAAGTTGGT GCAGGCTTAT CAACCAACAA 540

AGGTCATTCT TATCATGGCA GTGCCTTTGG TAAAGCACAG ACAGGATTTG GCCAGGTAGA 600

TGCCCTTGTC TCTTATAATC AAGTAAATGA CAGCGACTAT AAAGGCGGTA AAGGGTACAC 660

CAATCTGTTG GGCAATGATG TGGTTACAAG AAGTGCCTTG GATAAATCCA GTTACCTTGT 720

CAAGGCAGGG CTAACTGCTG GCGATCATCG ATTTGTGGTC AGCCATCTAA ATGAAGTTCA 780

TAAAGGCATT CGTGGCGTGC GTGAAGAGTT TGACTTCGCC AATCGTGCCT TGACGCTAGA 840

TATAGAGAAA AATAAGAAAA AACGTACCGA AGCAGAGCTT CAGGCAGAGT TAGATGACAA 900

ATATGCAGGC AAGGGTTACA AACTTGGCAG TAAAACACCA GATGGTAAAA AGTATAATGT 960

GGTTGATGCC AATGGTAAGT TGGTGGATAA TTTAGTTGGG AACAACCCAA CTCAGCGTGA 1020

AACCTACCAG AAGTTAACCA ACCTTGAATG GACAGCTAAA AACCTTGGTT TTGCGAATGA 1080

AGTTACTGCC AATGTCTATA AGTTAGAACA TGGACGCAAC TCCTCTAGCG ATAAAGGTAA 1140

CAGCTATATT CTTCGTGATG TACCTAATAC CATCAATGAT AACGGTGATA GCCCATCAAA 1200

TATGCATGTT AGTGCCAAAG GGGCTAATAT TAATTTTGAT AAAGAATTTA ATCACGGTCT 1260

ATTAAAAGGC TTTGGCGTTG ACCATACTTT ATTAAAATAT GGCATCAACT ATCGCCATCA 1320

AGAAGCTGTA CCGCCTAGAG GTATTAGACC TGGTTTTCAA CACCAAGAAA AAACCGATGC 1380

TGGCATTTAT CTAGAAGCGG TTAACCAAAT CAATGACTTT ACCATCAATA CAGGCGTGCG 1440

TGTTGACCGT TTTGACTTTA AAGCCATGGA CGGTAAAAAG GTTGGAAAAA CCGACATCAA 1500

CCCAAGCTTT GGGGTGATTT ATGATGTCAA TCCTAATCTT AGCGTCAGCG GTAACCTAAT 1560

CTATGCCACT CGCAGCCCAC GCTTTGCTGA TGCTATCCTA AGCCGTGGCT TCCGTGATGG 1620

CGTTGTGAGT ATTGCTGATA ACGCAAAAGC AGAAAAAGCA CGCAATACCG AGATTGGTTT 1680

TAACTATAAT AATGGGCCAT ATACCGCCTT TGGCAGCTAT TTTTGGCAGC GTGTGGATAA 1740

TGCCAGAGCT ACTGCCGATG CTGTACAACA CCCCACAGTT ACAACAGCTA AGATTACCTA 1800

TCTTGGCAAC CAAGGTCATC AGACCAACCA CGGTTATGAG CTGGGCGTAG GCTATACCGA 1860

AGGTGCGTGG CGTGCGCGTG CTGGCGTTGC TCACAGCAAG CCAACCATGC ACAATGTCAA 1920

ATTTAAAGCC AACCCTGAAT ATGCCGTGCG TACAGGTCGT ACATGGACAG CAGATGTCGC 1980

CTATCGCCTG CCAAACCCCA GTGTAGAGCT TGGTGTGAGA CACACATTGG TTGAAGGGGT 2040

AGATGCCAAA GACACTTCTA TCCTTAGCGG TAAATTTGAT GATAAAGATG GTGCTATTCT 2100

TAACCGTGAA GGCTATAATG TCAGTGACAT CTATGCCAAC TGGAAGCCTT ATGGCAATGA 2160

TAAGGTGAAT GTAAACTTTG CGGTGAATAA TGTCTTTAAT AAAAACTATC GCCCACACAC 2220

TCAGCGTGCT TCCATAGATA CCTTACCTGG GGCAGGTCGT GATTTCCGTG TTGGCGTGAA 2280

CTTCACTTAC TAATACTTGC CGATTTATCG GTATAATACT GAACACTCAA GCACGCTTGG 2340

GTGTTCTTTT TATGGGTATG AGTGGATAAA AACGATAAAA AAAGCCAATC GTATCATATT 2400

GATTGGCTAT AATGATAAAA TTAAATCATT ACTGATTTAA AAACCACGCA TTGGC 2455

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

Lys Tyr Ala Gly Lys
1 5

(2) INFORMATION FOR SEQ ID NO: 36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

Lys Tyr Ala Gly Lys Gly
1 5

(2) INFORMATION FOR SEQ ID NO: 37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

Lys Tyr Ala Gly Lys Gly Tyr
1 5

(2) INFORMATION FOR SEQ ID NO: 38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

Asn Lys Tyr Ala Gly Lys
1 5

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

Asn Lys Tyr Ala Gly Lys Gly
1 5

(2) INFORMATION FOR SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

Asn Lys Tyr Ala Gly Lys Gly Tyr
1 5

(2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

Leu Asp Ile Glu Lys Asp Lys Lys Lys Arg Thr Asp Glu Gln Leu Gln
1 5 10 15

Ala Glu Leu Asp Asn Lys Tyr Ala Gly Lys Gly Tyr
20 25

(2) INFORMATION FOR SEQ ID NO: 42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

Leu Asp Ile Lys Lys Asp Asp Lys Thr Leu Thr Glu Thr Glu Leu Gln
1 5 10 15

Ala Glu Leu Asp Asn Lys Tyr Ala Gly Lys Gly Tyr
20 25

(2) INFORMATION FOR SEQ ID NO: 43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

Leu Asp Ile Glu Lys Asp Lys Lys Lys Arg Thr Asp Glu Gln Leu Gln
1 5 10 15

Ala Glu Leu Asp Asp Lys Tyr Ala Gly Lys Gly Tyr
20 25

(2) INFORMATION FOR SEQ ID NO: 44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

Leu Asp Ile Glu Lys Asn Lys Lys Lys Arg Thr Glu Ala Glu Leu Gln
1 5 10 15

Ala Glu Leu Asp Asp Lys Tyr Ala Gly Lys Gly Tyr
20 25

(2) INFORMATION FOR SEQ ID NO: 45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

Ile Asp Ile Glu Lys Lys Gly Lys Ile Arg Thr Glu Ala Glu Leu Leu
1 5 10 15

Ala Glu Leu Asn Lys Asp Tyr Pro Gly Gln Gly Tyr
20 25

CLAIMS

1. An isolated peptide of about 5 to about 60 amino acid residues comprising at least the consecutive residues KYAGK (SEQ ID NO:35), wherein the isolated peptide is reactive with the antibody 10F3.

2. The isolated peptide of claim 1, wherein the consecutivc residues comprise at least KYAGKG (SEQ ID NO:36).

3. The isolated peptide of claim 2, wherein the consecutivc residues comprise at least KYAGKGY (SEQ ID NO:37).

4. The isolated peptide of claim 1, wherein the consecutivc residues comprise at least NKYAGK (SEQ ID NO:38).

5. The isolated peptide of claim 4, wherein the consecutivc residues comprise at least NKYAGKG (SEQ ID NO:39).

6. The isolated peptide of claim 5, wherein the consecutivc residues comprise at least NKYAGKGY (SEQ ID NO:40).

7. The isolated peptide of claim 1, wherein said peptide comprises at least about 20 consecutive residues of the amino acid sequence LDIEKDKKKRTDEQLQAELDNKYAGKGY (SEQ ID NO:41) or LDIKKDDKTLTETELQAELDNKYAGKGY (SEQ ID NO:42).

8. The isolated peptide of claim 7, wherein said peptide comprises at least about 20 consecutive residues of the amino acid sequence LDIEKDKKKRTDEQLQAELDNKYAGKGY (SEQ ID NO:41).

9. The isolated peptide of claim 7, wherein said peptide comprises at least about 20 consecutive residues of the amino acid sequence LDIKKDDKTLTETELQAELDNKYAGKGY (SEQ ID NO:42).

10. The isolated peptide of claim 1, wherein said peptide is about 5 amino acids in length.

11. The isolated peptide of claim 1, wherein said peptide is about 7 amino acids in length.

12. The isolated peptide of claim 1, wherein said peptide is about 10 amino acids in length.

13. The isolated peptide of claim 1, wherein said peptide is about 15 amino acids in length.

14. The isolated peptide of claim 1, wherein said peptide is about 20 amino acids in length.

15. The isolated peptide of claim 1, wherein said peptide is about 30 amino acids in length.

16. The isolated peptide of claim 1, wherein said peptide is about 40 amino acids in length.

17. The isolated peptide of claim 1, wherein said peptide is about 50 amino acids in length.

18. The isolated peptide of claim 1, wherein said peptide is about 60 amino acids in length.

19. An isolated peptide of about 20 to about 60 amino acids comprising at least about 20 consecutive residues of the amino acid sequence LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43), LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44) or IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45).

20. The isolated peptide of claim 19, wherein said peptide comprises at least about 20 consecutive residues of the amino acid sequence LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43).

21. The isolated peptide of claim 19, wherein said peptide comprises at least about 20 consecutive residues of the amino acid sequence LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44).

22. The isolated peptide of claim 19, wherein said peptide comprises at least about 20 consecutive residues of the amino acid sequence IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45).

23. The isolated peptide of claim 19, wherein said peptide is about 20 amino acids in length.

24. The isolated peptide of claim 19, wherein said peptide is about 30 amino acids in length.

25. The isolated peptide of claim 19, wherein said peptide is about 40 amino acids in length.

26. The isolated peptide of claim 19, wherein said peptide is about 50 amino acids in length.

27. The isolated peptide of claim 19, wherein said peptide is about 60 amino acids in length.

28. An antigenic composition comprising (a) an isolated peptide of about 5 to about 60 amino acids comprising the amino acid sequence KYAGK (SEQ ID NO:35) and (b) a pharmaceutically acceptable buffer or diluent.

29. The antigenic composition of claim 28, wherein said antigenic composition further comprises a carrier conjugated to said peptide.

30. The antigenic composition of claim 29, wherein said carrier is KLH.

31. The antigenic composition of claim 28, further comprising an adjuvant.

32. The antigenic composition of claim 31, wherein said adjuvant comprises a lipid.

33. The antigenic composition of claim 28, wherein said peptide is covalently linked to a second antigen.

34. The antigenic composition of claim 33, wherein said second antigen is a peptide antigen.

35. The antigenic composition of claim 34, wherein said second antigen is a non-peptide antigen.

36. An antigenic composition comprising (a) an isolated peptide of about 20 to about 60 amino acids comprising the amino acid sequence LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43), LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44) or IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45) and (b) a pharmaceutically acceptable buffer or diluent.

37. The antigenic composition of claim 36, further comprising a carrier conjugated to said peptide.

38. The antigenic composition of claim 37, wherein said carrier is KLH.

39. The antigenic composition of claim 36, further comprising an adjuvant.

40. The antigenic composition of claim 39, wherein said adjuvant comprises a lipid.

41. The antigenic composition of claim 36, wherein said peptide is covalently linked to a second antigen.

42. The antigenic composition of claim 41, wherein said second antigen is a peptide antigen.

43. The antigenic composition of claim 41, wherein said second antigen is a non-peptide antigen.

44. A vaccine composition comprising an isolated peptide of about 20 to about 60 amino acids comprising the amino acid sequence KYAGK (SEQ ID NO:35) and a pharmaceutically acceptable buffer or diluent.

45. A vaccine composition comprising an isolated peptide of about 20 to about 60 amino acids comprising the amino acid sequence LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43), LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44) or IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45) and a pharmaceutically acceptable buffer or diluent.

46. A vaccine composition of claim 45, wherein said vaccine composition comprises at least about 20 consecutive residues of the amino acid sequence LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43).

47. A vaccine composition of claim 45, wherein said vaccine composition comprises at least about 20 consecutive residues of the amino acid sequence LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44).

48. A vaccine composition of claim 45, wherein said vaccine composition comprises at least about 20 consecutive residues of the amino acid sequence IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45).

49. A method for inducing an immune response in a mammal comprising the step of providing to said mammal an antigenic composition comprising (a) an isolated peptide of about 5 to about 60 amino acids comprising the amino acid sequence KYAGK (SEQ ID NO:35) and (b) a pharmaceutically acceptable buffer or diluent.

50. A method for inducing an immune response in a mammal comprising the step of providing to said mammal an antigenic composition comprising (a) an isolated peptide of about 20 to about 60 amino acids comprising the amino acid sequence LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43), LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44) or IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45) and (b) a pharmaceutically acceptable buffer or diluent.

51. The isolated peptide of claim 50, wherein said peptide comprises at least about 20 consecutive residues of the amino acid sequence LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43).

52. The isolated peptide of claim 50, wherein said peptide comprises at least about 20 consecutive residues of the amino acid sequence LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44).

53. The isolated peptide of claim 50, wherein said peptide, comprises at least about 20 consecutive residues of the amino acid sequence IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45).

54. An isolated and purified *M. catarrhalis* CopB antigen comprising the amino acid sequence sequence KYAGK (SEQ ID NO:35).

55. An isolated and purified *M. catarrhalis* CopB antigen comprising the amino acid sequence LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43), LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44) or IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45).

56. The CopB antigen of claim 55, wherein said antigen comprises at least about 20 consecutive residues of the amino acid sequence LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43).

57. The CopB antigen of claim 55, wherein said antigen comprises at least about 20 consecutive residues of the amino acid sequence LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44).

58. The CopB antigen of claim 55, wherein said antigen comprises at least about 20 consecutive residues of the amino acid sequence IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45).

59. A nucleic acid encoding the CopB antigen of the *M. catarrhalis* isolate TTA24.

60. A nucleic acid having the *copB* DNA sequence, SEQ ID NO:34, of the *M. catarrhalis* isolate TTA24.

61. A nucleic acid encoding the CopB antigen of the *M. catarrhalis* isolate O12E.

62. A nucleic acid having the *copB* DNA sequence, SEQ ID NO:28, of the *M. catarrhalis* isolate O12E.

63. A nucleic acid encoding the Cop B antigen of the *M. catarrhalis* isolate O46E.

64. A nucleic acid having the *copB* DNA sequence, SEQ ID NO:32, of the *M. catarrhalis* isolate O46E.

65. A method for generating a strain-specific antibody to *M. catarrhalis* CopB antigen comprising the step of providing to a mammal an antigenic composition comprising (a) an isolated peptide comprising an amino acid sequence corresponding to residues of variable regions of said CopB antigen and (b) a pharmaceutically acceptable buffer or diluent.

66. A method for diagnosing *M. catarrhalis* infection comprising the step of determining the presence, in a sample, of an *M. catarrhalis,* amino acid sequence corresponding to residues of variable regions of said CopB antigen.

67. The method of claim 66, wherein said determining comprises PCR.

68. The method of claim 66, wherein said determining comprises immunologic reactivity of an antibody with an *M. catarrhalis* antigen.

69. A method for generating a strain-common antibody to *M. catarrhalis* CopB antigen comprising the step of providing to a mammal an antigenic composition comprising (a) an isolated peptide comprising an amino acid sequence corresponding to residues of common regions of said CopB antigen and (b) a pharmaceutically acceptable buffer or diluent.

70. A method for diagnosing *M. catarrhalis* infection comprising the step of determining the presence, in a sample, of an *M. catarrhalis* amino acid sequence corresponding to residues of common regions of said CopB antigen.

71. The method of claim 70, wherein said determining comprises PCR.

72. The method of claim 70, wherein said determining comprises immunologic reactivity of an antibody with an *M. catarrhalis* antigen.

73. A method for treating an individual having an *M. catarrhalis* infection comprising providing to said individual an isolated peptide of about 20 to about 60 amino acids comprising at least the consecutive residues KYAGK (SEQ ID NO:35).

74. A method for preventing or limiting an *M. catarrhalis* infection comprising providing to a subject an isolated peptide of about 20 to about 60 amino acids comprising at least about 20 consecutive residues of the amino acid sequence LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43), LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44) or IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45).

75. The peptide of claim 74, wherein the said peptide comprises at least about 20 consecutive residues of the amino acid sequence LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43).

76. The peptide of claim 74, wherein the said peptide comprises at least about 20 consecutive residues of the amino acid sequence LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44).

77. The peptide of claim 74, wherein the said peptide comprises at least about 20 consecutive residues of the amino acid sequence IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45).

78. A method for preventing or limiting an *M. catarrhalis* infection comprising providing to a subject an antibody that reacts immunologically with an epitope formed by the amino acid sequence KYAGK (SEQ ID NO:35).

79. A method for preventing or limiting an *M. catarrhalis* infection comprising providing to a subject an antibody that reacts immunologically with an epitope formed by the amino acid sequence LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43), LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44) or IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45).

80. The epitope of claim 79, wherein said epitope comprises at least about 20 consecutive residues of the amino acid sequence sequence LDIEKDKKKRTDEQLQAELDDKYAGKGY (SEQ ID NO:43).

81. The epitope of claim 79, wherein said epitope comprises at least about 20 consecutive residues of the amino acid sequence sequence LDIEKNKKKRTEAELQAELDDKYAGKGY (SEQ ID NO:44).

82. The epitope of claim 79, wherein said epitope comprises at least about 20 consecutive residues of the amino acid sequence sequence IDIEKKGKIRTEAELLAELNKDYPGQGY (SEQ ID NO:45).

83. A method for screening a peptide for reactivity with a CopB antibody comprising the steps of:

(a) providing said peptide;
(b) contacting said peptide with an antibody that binds immunogically to CopB; and
(c) determining the binding of said antibody to said peptide.

84. The method of claim 83, wherein said antibody is 10F3.

85. The method of claim 83, wherein said determining comprises an immunoassay selected from the group consisting of a Western blot, an ELISA, an RIA and immunoaffinity separation.

86. A method for screening a peptide for the ability to induce a protective immune response against *M. catarrhalis* comprising the steps of:

(a) providing said peptide;
(b) administering a peptide in a suitable form to an experimental animal; (c) challenging said animal with *M. catarrhalis;* and
(c) assaying the infection of said animal with *M. catarrhalis.*

87. The method of claim 86, wherein said animal is a mouse, said challenging is a pulmonary challenge, and said assaying comprises assessing the degree of pulmonary clearance by said mouse.

88. The method of claim 86, wherein said peptide is a CopB peptide.

89. The method of claim 88, wherein said CopB peptide encompasses residues 296-300 of *M. catarrhalis* strain 035E.

90. The method of claim 89, wherein said CopB peptide is a peptide having at least 6 consecutive amino acids from *M. catarrhalis* strain 035E, including residue 295.

91. An isolated peptide having at least about 5 consecutive amino acids from the CopB protein of *M. catarrhalis,* wherein said peptide includes residue 296-300, or the analogous positions thereof when compared to strain 035E.

92. The isolated peptide of claim 91 wherein said CopB peptide is a peptide having at least 6 consecutive amino acids and includes residue 295, or the analogous position thereof when compared to strain 035E.

93. The isolated peptide of claim 91, wherein said peptide is between 6 and 60 amino acids in length.

94. The isolated peptide of claim 93, wherein said peptide comprises non-CopB sequences.

95. The isolated peptide of claim 94, wherein said peptide comprises non-*M. catarrhalis* sequences.

96. An antigen composition comprising (i) an isolated peptide having at least about 6 consecutive amino acids from the CopB protein of *M. catarrhalis*, wherein said peptide includes residue 295, or the analogous position thereof when compared to strain 035E and (ii) a pharmaceutically acceptable buffer or diluent.

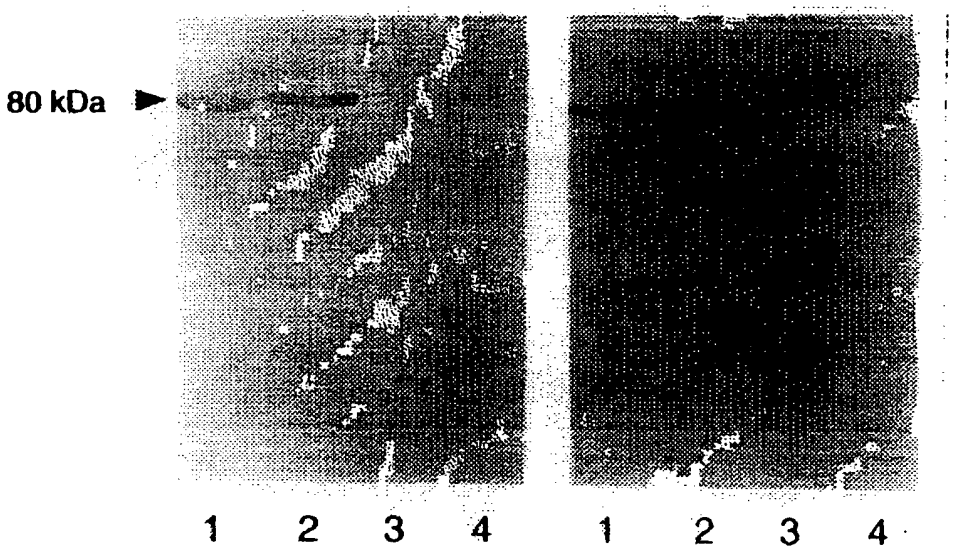


FIG. 1A FIG. 1B

Region 1 of native CopB proteins

	275 280 290 300		Reactivity with Antibody Preparations in Western Blot		
			MAb 10F3	GST-26 antiserum	GST antiserum
035E	L D I E K D K K K R T D E Q L Q A E L D N K Y A G K	(SEQ. ID NO:1)	+	+	-
012E	L D I K K D D K T L T E T E L Q A E L D N K Y A G K	(SEQ. ID NO:2)	+	-	-
TTA24	L D I E K N K K K R T E A E L Q A E L D D K Y A G K	(SEQ. ID NO:4)	-	-	-
046E	I D I E K K G K I R T E A E L L A E L N K D Y P G Q	(SEQ. ID NO:5)	-	-	-

Synthetic peptides

	1 10 20 26		Reactivity with Antibody Preparations in ELISA		
			MAb 10F3*	GST-26 antiserum	GST antiserum
R1	L D I E K D K K K R T D E Q L Q A E L D N K Y A G K	(SEQ. ID NO:1)	1.670	1.624	0.066
R1A	L D I E K N K K K R T D E Q L Q A E L D N K Y A G K	(SEQ. ID NO:7)	1.440	1.402	0.038
R1E	D E Q L Q A E L D N K Y A G K	(SEQ. ID NO:6)	0.043	0.726	0.068
R1B	L D I E K D K K K R T D E Q L Q A E L D D K Y A G K	(SEQ. ID NO:3)	0.024	1.141	0.044
R1-TTA24	L D I E K N K K K R T E A E L Q A E L D D K Y A G K	(SEQ. ID NO:4)	0.003	1.278	0.040

FIG. 2

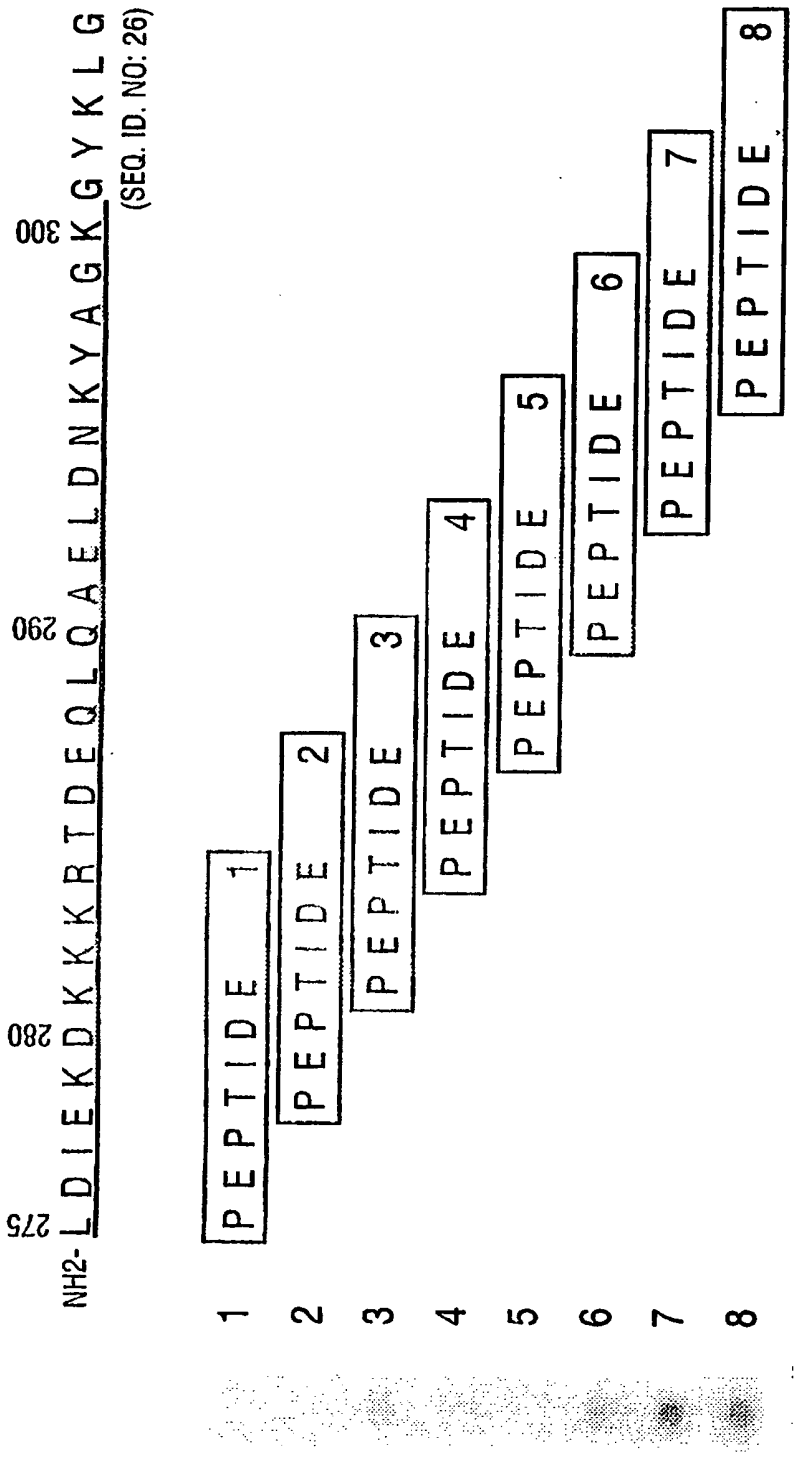

NH2-LDIEKDKKKRTDEQLQAELDNKYAGKGYKLG
(SEQ. ID. NO: 26)
275
280
290
300
PEPTIDE 1
PEPTIDE 2
PEPTIDE 3
PEPTIDE 4
PEPTIDE 5
PEPTIDE 6
PEPTIDE 7
PEPTIDE 8
1
2
3
4
5
6
7
8

FIG. 3

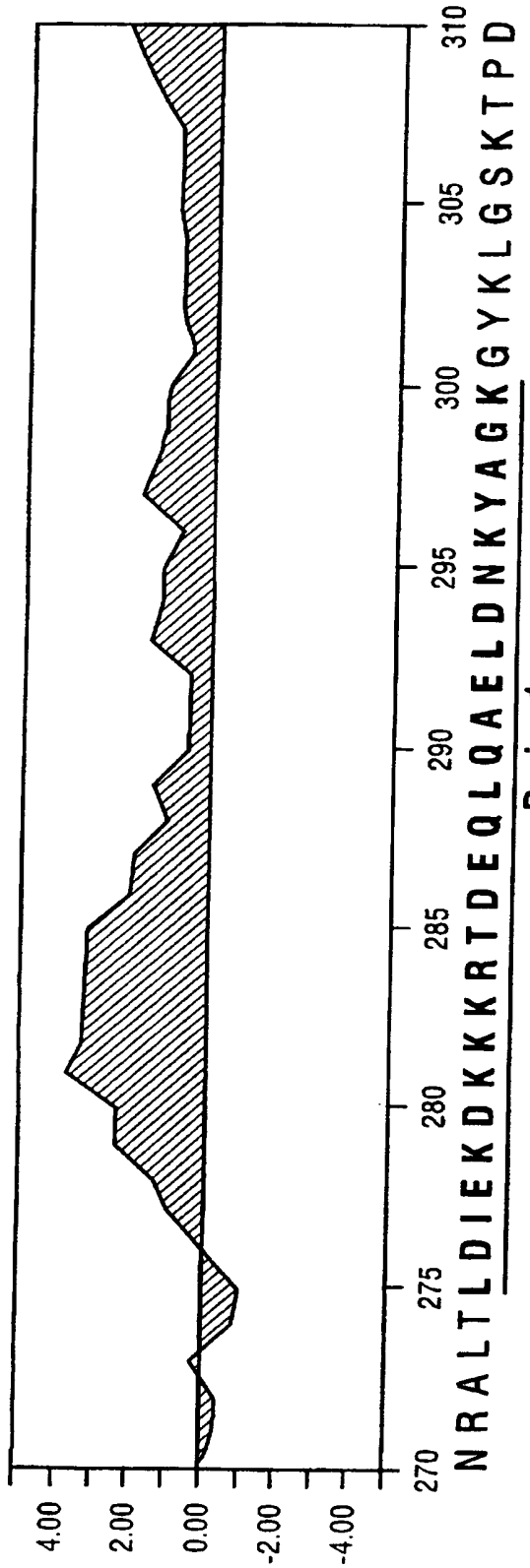

4.00
2.00
0.00
-2.00
-4.00
270
275
280
285
290
295
300
305
310
N R A L T L D I E K D K K K R T D E Q L Q A E L D N K Y A G K G Y K L G S K T P D
Region 1

FIG. 4A

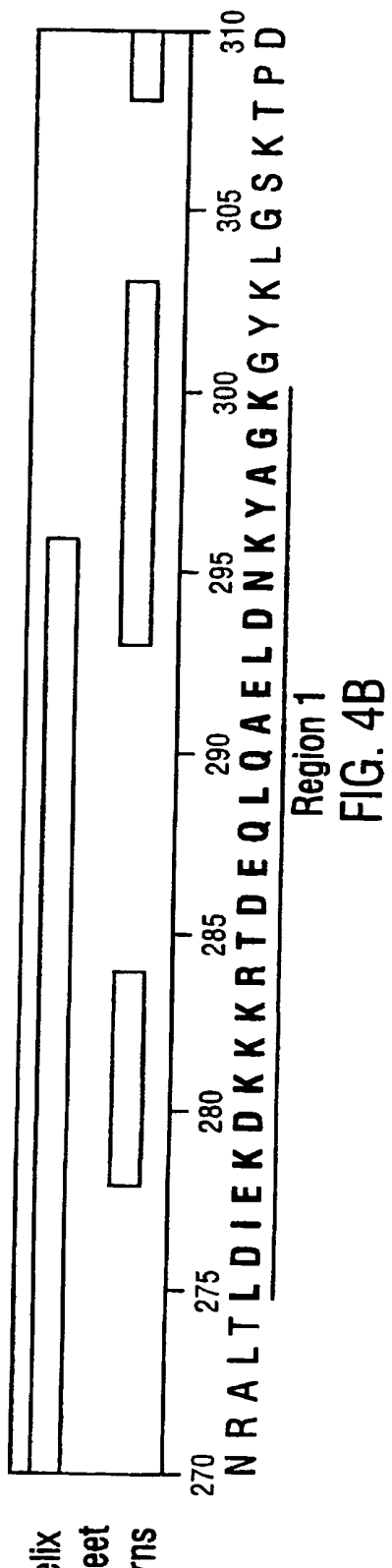

Helix
Sheet
Turns
270
275
280
285
290
295
300
305
310
N R A L T L D I E K D K K K R T D E Q L Q A E L D N K Y A G K G Y K L G S K T P D
Region 1

FIG. 4B

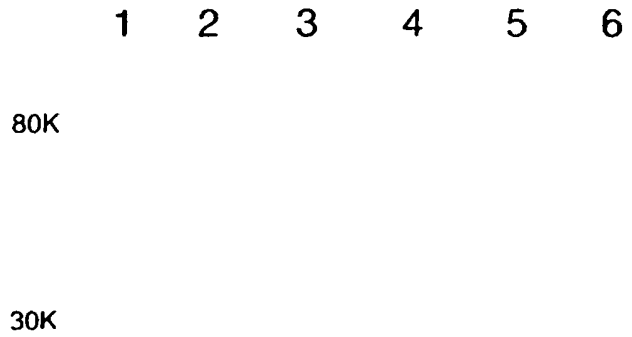


FIG. 5A

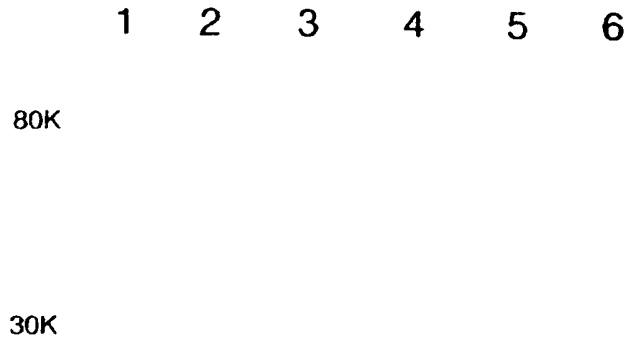


FIG. 5B

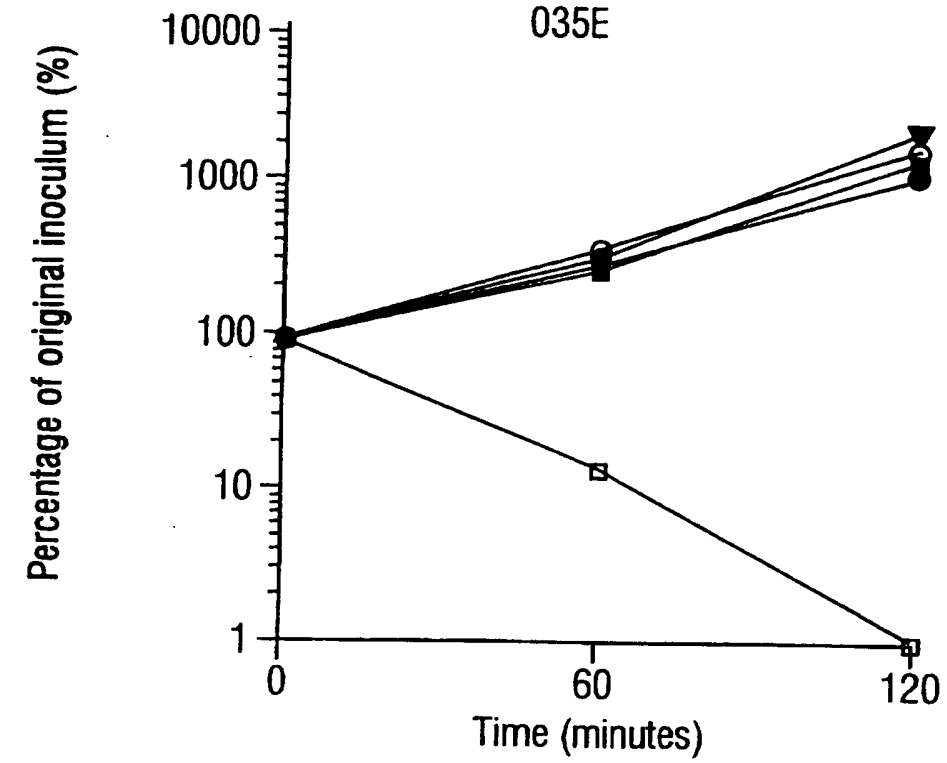

035E
10000
1000
100
10
1
Percentage of original inoculum (%)
0
60
120
Time (minutes)

FIG. 6A

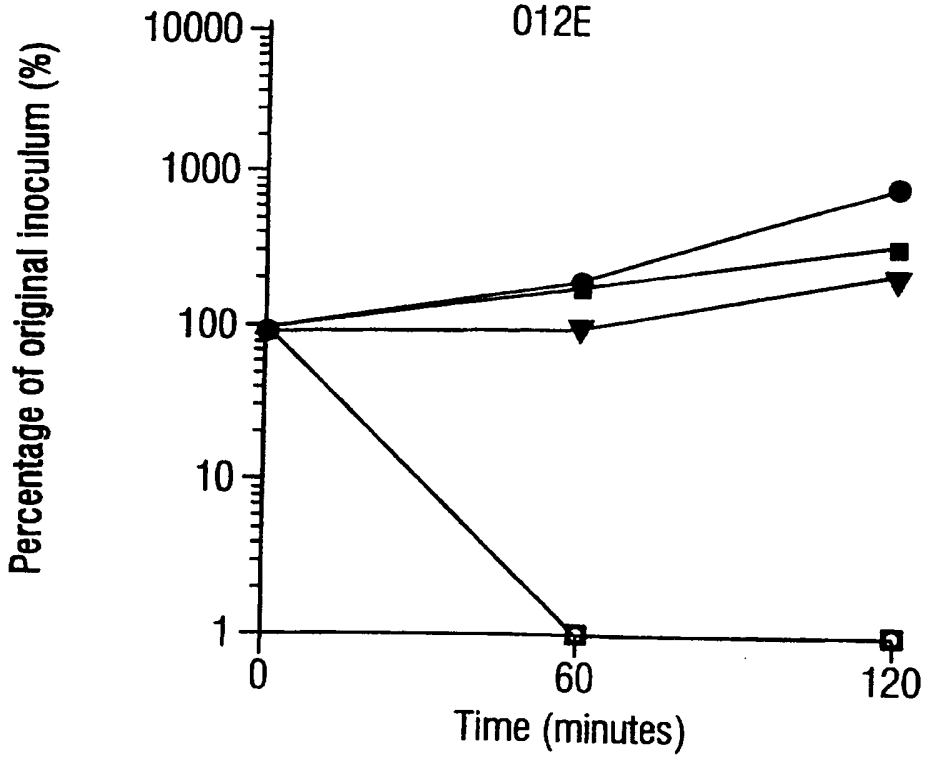

012E
10000
1000
100
10
1
Percentage of original inoculum (%)
0
60
120
Time (minutes)

FIG. 6B

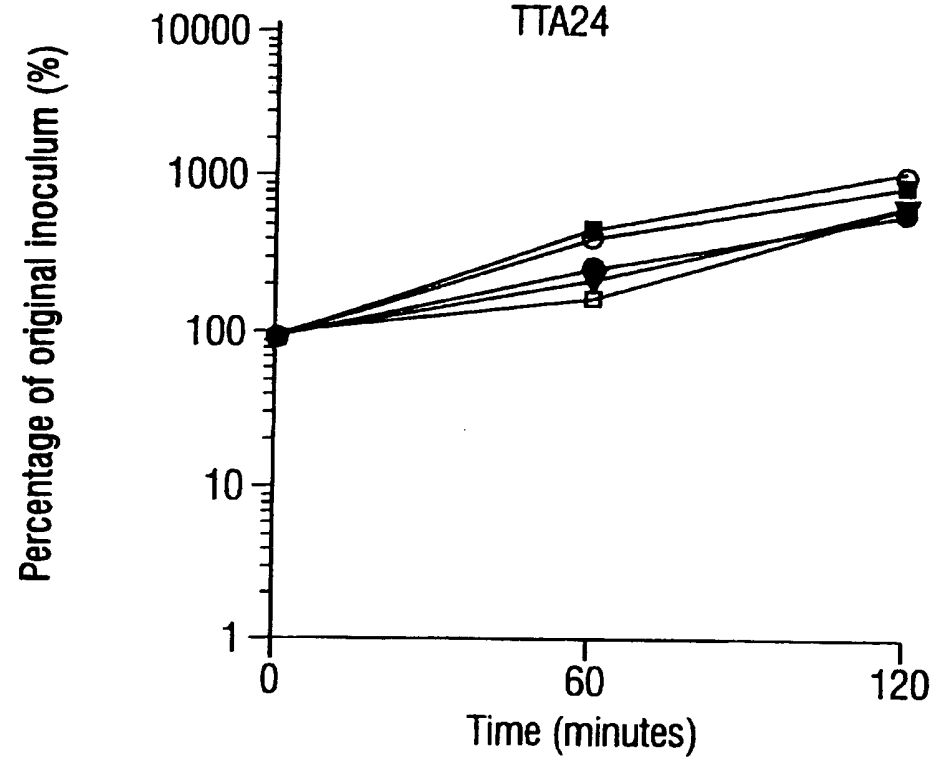

TTA24
10000
1000
100
10
1
Percentage of original inoculum (%)
0
60
120
Time (minutes)

FIG. 6C

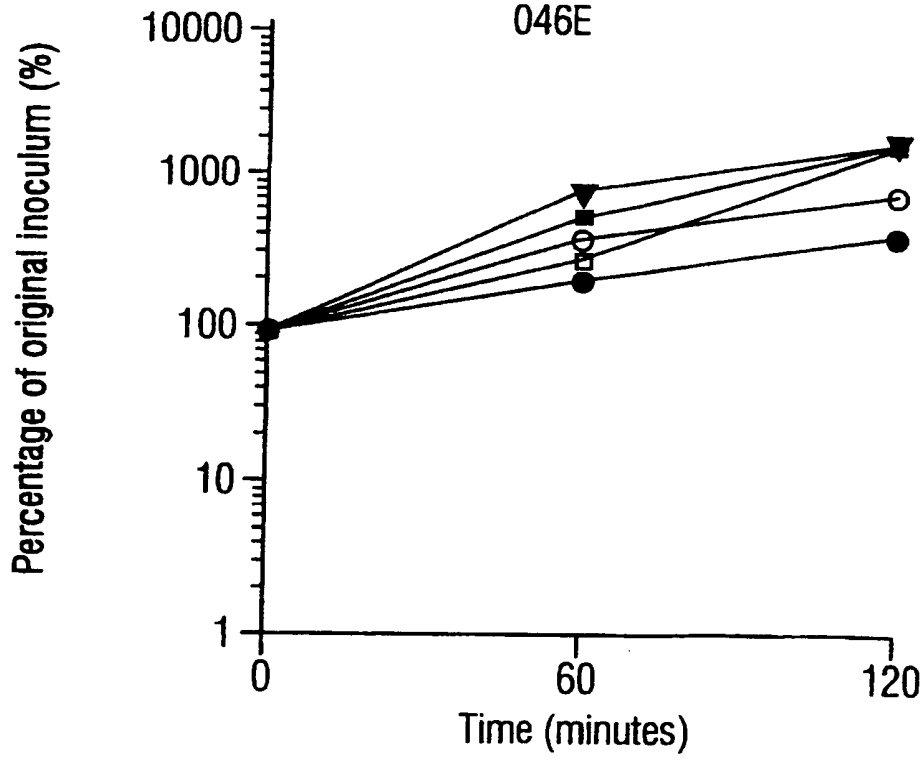

046E
10000
1000
100
10
1
Percentage of original inoculum (%)
0
60
120
Time (minutes)

FIG. 6D